(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 September 2001 (27.09.2001)

PCT

(10) International Publication Number
WO 01/70976 A2

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, C12N 15/62, C07K 16/18, C12Q 1/68, G01N 33/53, A61K 39/00, 39/395

(21) International Application Number: PCT/US01/09062

(22) International Filing Date: 20 March 2001 (20.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/190,710	21 March 2000 (21.03.2000)	US
60/213,748	22 June 2000 (22.06.2000)	US
60/257,276	19 December 2000 (19.12.2000)	US

(71) Applicant *(for all designated States except US)*: **CORIXA CORPORATION** [US/US]; Suite 200, 1124 Columbia Street, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **XU, Jiangchun** [US/US]; 15805 S.E. 43rd Place, Bellevue, WA 98006 (US). **PYLE, Ruth, A.** [US/US]; 100 N.E. 62nd Street, Seattle, WA 98115 (US). **STOLK, John, A.** [US/US]; 7436 Northeast 144th Place, Bothell, WA 98011 (US).

(74) Agents: **POTTER, Jane, E., R.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/70976 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, such as ovarian or endometrial cancer, are disclosed. Compositions may comprise one or more ovarian carcinoma proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses such an antigen, or a T cell that is specific for cells expressing such an antigen. Such compositions may be used, for example, for the prevention and treatment of diseases such as ovarian and endometrial cancer. Diagnostic methods based on detecting an ovarian carcinoma protein, or mRNA encoding such an antigen, in a sample are also provided.

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, such as ovarian or endometrial cancer. The invention is more specifically related to polypeptides comprising at least a portion of an ovarian carcinoma protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of cancers such as ovarian and endometrial cancer, and for the diagnosis and monitoring of such cancers.

BACKGROUND OF THE INVENTION

Ovarian cancer is a significant health problem for women in the United States and throughout the world. Although advances have been made in detection and therapy of this cancer, no vaccine or other universally successful method for prevention or treatment is currently available. Management of the disease currently relies on a combination of early diagnosis and aggressive treatment, which may include one or more of a variety of treatments such as surgery, radiotherapy, chemotherapy and hormone therapy. The course of treatment for a particular cancer is often selected based on a variety of prognostic parameters, including an analysis of specific tumor markers. However, the use of established markers often leads to a result that is difficult to interpret, and high mortality continues to be observed in many cancer patients.

Immunotherapies have the potential to substantially improve cancer treatment and survival. Such therapies may involve the generation or enhancement of an immune response to an ovarian carcinoma protein. However, to date, relatively few ovarian carcinoma proteins are known and the generation of an immune response against such antigens has not been shown to be therapeutically beneficial.

Accordingly, there is a need in the art for improved methods for detecting and treating cancers such as ovarian cancer. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as ovarian and endometrial cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of an ovarian carcinoma protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in SEQ ID NO:1-222; (b) variants of a sequence recited in SEQ ID NO:1-222 and (c) complements of a sequence of (a) or (b).

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of an ovarian carcinoma protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, vaccines for prophylactic or therapeutic use are provided. Such vaccines comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to an ovarian carcinoma protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above. The patient may be afflicted with ovarian or endometrial cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with an ovarian carcinoma protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for an ovarian carcinoma protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating $CD4^+$ and/or $CD8^+$ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of an ovarian carcinoma protein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expresses such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be ovarian or endometrial cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a)

contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes an ovarian carcinoma protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes an ovarian carcinoma protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE SEQUENCE IDENTIFIERS

SEQ ID NOs: 1-41 are identified in Example 1.

SEQ ID NO:42 is the determined cDNA sequence for clone R0198:A03
SEQ ID NO:43 is the determined cDNA sequence for clone R0198:A07
SEQ ID NO:44 is the determined cDNA sequence for clone R0198:A08
SEQ ID NO:45 is the determined cDNA sequence for clone R0198:A09
SEQ ID NO:46 is the determined cDNA sequence for clone R0198:B01
SEQ ID NO:47 is the determined cDNA sequence for clone R0198:B02
SEQ ID NO:48 is the determined cDNA sequence for clone R0198:B04
SEQ ID NO:49 is the determined cDNA sequence for clone R0198:B08
SEQ ID NO:50 is the determined cDNA sequence for clone R0198:B11
SEQ ID NO:51 is the determined cDNA sequence for clone R0198:C01
SEQ ID NO:52 is the determined cDNA sequence for clone R0198:C02
SEQ ID NO:53 is the determined cDNA sequence for clone R0198:C03
SEQ ID NO:54 is the determined cDNA sequence for clone R0198:C04
SEQ ID NO:55 is the determined cDNA sequence for clone R0198:C05
SEQ ID NO:56 is the determined cDNA sequence for clone R0198:C06
SEQ ID NO:57 is the determined cDNA sequence for clone R0198:C08
SEQ ID NO:58 is the determined cDNA sequence for clone R0198:C09
SEQ ID NO:59 is the determined cDNA sequence for clone R0198:C10
SEQ ID NO:60 is the determined cDNA sequence for clone R0198:C12
SEQ ID NO:61 is the determined cDNA sequence for clone R0198:D01
SEQ ID NO:62 is the determined cDNA sequence for clone R0198:D02
SEQ ID NO:63 is the determined cDNA sequence for clone R0198:D03
SEQ ID NO:64 is the determined cDNA sequence for clone R0198:D04
SEQ ID NO:65 is the determined cDNA sequence for clone R0198:D05
SEQ ID NO:66 is the determined cDNA sequence for clone R0198:D06
SEQ ID NO:67 is the determined cDNA sequence for clone R0198:D07
SEQ ID NO:68 is the determined cDNA sequence for clone R0198:D08
SEQ ID NO:69 is the determined cDNA sequence for clone R0198:D09
SEQ ID NO:70 is the determined cDNA sequence for clone R0198:D11
SEQ ID NO:71 is the determined cDNA sequence for clone R0198:E01
SEQ ID NO:72 is the determined cDNA sequence for clone R0198:E03
SEQ ID NO:73 is the determined cDNA sequence for clone R0198:E05

SEQ ID NO:74 is the determined cDNA sequence for clone R0198:E06

SEQ ID NO:75 is the determined cDNA sequence for clone R0198:E09

SEQ ID NO:76 is the determined cDNA sequence for clone R0198:E10

SEQ ID NO:77 is the determined cDNA sequence for clone R0198:E11

SEQ ID NO:78 is the determined cDNA sequence for clone R0198:E12

SEQ ID NO:79 is the determined cDNA sequence for clone R0198:F01

SEQ ID NO:80 is the determined cDNA sequence for clone R0198:F02

SEQ ID NO:81 is the determined cDNA sequence for clone R0198:F03

SEQ ID NO:82 is the determined cDNA sequence for clone R0198:F04

SEQ ID NO:83 is the determined cDNA sequence for clone R0198:F06

SEQ ID NO:84 is the determined cDNA sequence for clone R0198:F07

SEQ ID NO:85 is the determined cDNA sequence for clone R0198:F09

SEQ ID NO:86 is the determined cDNA sequence for clone R0198:F10

SEQ ID NO:87 is the determined cDNA sequence for clone R0198:F11

SEQ ID NO:88 is the determined cDNA sequence for clone R0198:F12

SEQ ID NO:89 is the determined cDNA sequence for clone R0198:G01

SEQ ID NO:90 is the determined cDNA sequence for clone R0198:G02

SEQ ID NO:91 is the determined cDNA sequence for clone R0198:G03

SEQ ID NO:92 is the determined cDNA sequence for clone R0198:G04

SEQ ID NO:93 is the determined cDNA sequence for clone R0198:G05

SEQ ID NO:94 is the determined cDNA sequence for clone R0198:G06

SEQ ID NO:95 is the determined cDNA sequence for clone R0198:G09

SEQ ID NO:96 is the determined cDNA sequence for clone R0198:G11

SEQ ID NO:97 is the determined cDNA sequence for clone R0198:G12

SEQ ID NO:98 is the determined cDNA sequence for clone R0198:H01

SEQ ID NO:99 is the determined cDNA sequence for clone R0198:H03

SEQ ID NO:100 is the determined cDNA sequence for clone R0198:H04

SEQ ID NO:101 is the determined cDNA sequence for clone R0198:H06

SEQ ID NO:102 is the determined cDNA sequence for clone R0198:H09

SEQ ID NO:103 is the determined cDNA sequence for clone R0198:H10

SEQ ID NO:104 is the determined cDNA sequence for clone R0199:A03

SEQ ID NO:105 is the determined cDNA sequence for clone R0199:A05

SEQ ID NO:106 is the determined cDNA sequence for clone R0199:A06
SEQ ID NO:107 is the determined cDNA sequence for clone R0199:A07
SEQ ID NO:108 is the determined cDNA sequence for clone R0199:A08
SEQ ID NO:109 is the determined cDNA sequence for clone R0199:A11
SEQ ID NO:110 is the determined cDNA sequence for clone R0199:B01
SEQ ID NO:111 is the determined cDNA sequence for clone R0199:B03
SEQ ID NO:112 is the determined cDNA sequence for clone R0199:B06
SEQ ID NO:113 is the determined cDNA sequence for clone R0199:B07
SEQ ID NO:114 is the determined cDNA sequence for clone R0199:B08
SEQ ID NO:115 is the determined cDNA sequence for clone R0199:B09
SEQ ID NO:116 is the determined cDNA sequence for clone R0199:B11
SEQ ID NO:117 is the determined cDNA sequence for clone R0199:C01
SEQ ID NO:118 is the determined cDNA sequence for clone R0199:C02
SEQ ID NO:119 is the determined cDNA sequence for clone R0199:C06
SEQ ID NO:120 is the determined cDNA sequence for clone R0199:C07
SEQ ID NO:121 is the determined cDNA sequence for clone R0199:C08
SEQ ID NO:122 is the determined cDNA sequence for clone R0199:C09
SEQ ID NO:123 is the determined cDNA sequence for clone R0199:C10
SEQ ID NO:124 is the determined cDNA sequence for clone R0199:C11
SEQ ID NO:125 is the determined cDNA sequence for clone R0199:C12
SEQ ID NO:126 is the determined cDNA sequence for clone R0199:D01
SEQ ID NO:127 is the determined cDNA sequence for clone R0199:D02
SEQ ID NO:128 is the determined cDNA sequence for clone R0199:D04
SEQ ID NO:129 is the determined cDNA sequence for clone R0199:D06
SEQ ID NO:130 is the determined cDNA sequence for clone R0199:D07
SEQ ID NO:131 is the determined cDNA sequence for clone R0199:D08
SEQ ID NO:132 is the determined cDNA sequence for clone R0199:D11
SEQ ID NO:133 is the determined cDNA sequence for clone R0199:E02
SEQ ID NO:134 is the determined cDNA sequence for clone R0199:E03
SEQ ID NO:135 is the determined cDNA sequence for clone R0199:E05
SEQ ID NO:136 is the determined cDNA sequence for clone R0199:E06
SEQ ID NO:137 is the determined cDNA sequence for clone R0199:E-8

SEQ ID NO:138 is the determined cDNA sequence for clone R0199:E09
SEQ ID NO:139 is the determined cDNA sequence for clone R0199:E10
SEQ ID NO:140 is the determined cDNA sequence for clone R0199:E12
SEQ ID NO:141 is the determined cDNA sequence for clone R0199:F01
SEQ ID NO:142 is the determined cDNA sequence for clone R0199:F03
SEQ ID NO:143 is the determined cDNA sequence for clone R0199:F04
SEQ ID NO:144 is the determined cDNA sequence for clone R0199:F06
SEQ ID NO:145 is the determined cDNA sequence for clone R0199:F09
SEQ ID NO:146 is the determined cDNA sequence for clone R0199:F10
SEQ ID NO:147 is the determined cDNA sequence for clone R0199:G01
SEQ ID NO:148 is the determined cDNA sequence for clone R0199:G05
SEQ ID NO:149 is the determined cDNA sequence for clone R0199:G06
SEQ ID NO:150 is the determined cDNA sequence for clone R0199:G08
SEQ ID NO:151 is the determined cDNA sequence for clone R0199:G11
SEQ ID NO:152 is the determined cDNA sequence for clone R0199:G12
SEQ ID NO:153 is the determined cDNA sequence for clone R0199:H02
SEQ ID NO:154 is the determined cDNA sequence for clone R0199:H03
SEQ ID NO:155 is the determined cDNA sequence for clone R0200:A05
SEQ ID NO:156 is the determined cDNA sequence for clone R0200:A06
SEQ ID NO:157 is the determined cDNA sequence for clone R0200:A10
SEQ ID NO:158 is the determined cDNA sequence for clone R0200:A12
SEQ ID NO:159 is the determined cDNA sequence for clone R0200:B03
SEQ ID NO:160 is the determined cDNA sequence for clone R0200:B04
SEQ ID NO:161 is the determined cDNA sequence for clone R0200:B07
SEQ ID NO:162 is the determined cDNA sequence for clone R0200:B08
SEQ ID NO:163 is the determined cDNA sequence for clone R0200:B12
SEQ ID NO:164 is the determined cDNA sequence for clone R0200:C02
SEQ ID NO:165 is the determined cDNA sequence for clone R0200:C07
SEQ ID NO:166 is the determined cDNA sequence for clone R0200:C09
SEQ ID NO:167 is the determined cDNA sequence for clone R0200:C10
SEQ ID NO:168 is the determined cDNA sequence for clone R0200:D01
SEQ ID NO:169 is the determined cDNA sequence for clone R0200:D03

SEQ ID NO:170 is the determined cDNA sequence for clone R0200:D05
SEQ ID NO:171 is the determined cDNA sequence for clone R0200:D06
SEQ ID NO:172 is the determined cDNA sequence for clone R0200:D07
SEQ ID NO:173 is the determined cDNA sequence for clone R0200:D08
SEQ ID NO:174 is the determined cDNA sequence for clone R0200:D09
SEQ ID NO:175 is the determined cDNA sequence for clone R0200:D11
SEQ ID NO:176 is the determined cDNA sequence for clone R0200:D12
SEQ ID NO:177 is the determined cDNA sequence for clone R0200:E03
SEQ ID NO:178 is the determined cDNA sequence for clone R0200:E04
SEQ ID NO:179 is the determined cDNA sequence for clone R0200:E06
SEQ ID NO:180 is the determined cDNA sequence for clone R0200:E07
SEQ ID NO:181 is the determined cDNA sequence for clone R0200:E08
SEQ ID NO:182 is the determined cDNA sequence for clone R0200:E09
SEQ ID NO:183 is the determined cDNA sequence for clone R0200:E12
SEQ ID NO:184 is the determined cDNA sequence for clone R0200:F01
SEQ ID NO:185 is the determined cDNA sequence for clone R0200:F04
SEQ ID NO:186 is the determined cDNA sequence for clone R0200:F05
SEQ ID NO:187 is the determined cDNA sequence for clone R0200:F07
SEQ ID NO:188 is the determined cDNA sequence for clone R0200:F08
SEQ ID NO:189 is the determined cDNA sequence for clone R0200:F09
SEQ ID NO:190 is the determined cDNA sequence for clone R0200:F10
SEQ ID NO:191 is the determined cDNA sequence for clone R0200:F11
SEQ ID NO:192 is the determined cDNA sequence for clone R0200:F12
SEQ ID NO:193 is the determined cDNA sequence for clone R0200:G02
SEQ ID NO:194 is the determined cDNA sequence for clone R0200:G07
SEQ ID NO:195 is the determined cDNA sequence for clone R0200:G08
SEQ ID NO:196 is the determined cDNA sequence for clone R0200:G09
SEQ ID NO:197 is the determined cDNA sequence for clone R0200:G10
SEQ ID NO:198 is the determined cDNA sequence for clone R0200:G12
SEQ ID NO:199 is the determined cDNA sequence for clone R0200:H03
SEQ ID NO:200 is the determined cDNA sequence for clone R0200:H05
SEQ ID NO:201 is the determined cDNA sequence for clone R0200:H07

SEQ ID NO:202 is the determined cDNA sequence for clone R0200:H09

SEQ ID NO:203 is the determined cDNA sequence for clone R0200:H11

SEQ ID NO:204 is the determined cDNA sequence for clone 57877.2

SEQ ID NO:205 is the determined cDNA sequence for clone57879.3

SEQ ID NO:206 is the determined cDNA sequence for clone57881.2

SEQ ID NO:207 is the determined cDNA sequence for clone57882.1

SEQ ID NO:208 is the determined cDNA sequence for clone57884.2

SEQ ID NO:209 is the determined cDNA sequence for clone57888.2

SEQ ID NO:210 is an extended cDNA sequence for clone RO198 C12 (SEQ ID NO: 60), also referred to as O593S

SEQ ID NO:211 is an extended cDNA sequence for clone RO198 F2 (SEQ ID NO: 80), also referred to as O594S

SEQ ID NO:212 is an extended cDNA sequence for clone RO199 A7 (SEQ ID NO: 107), also referred to as O595S

SEQ ID NO:213 is an extended cDNA sequence for clone RO199 C12 (SEQ ID NO: 125), also referred to as O596S

SEQ ID NO:214 is a full length cDNA sequence for HSPCO67, a sequence having homology with O596S

SEQ ID NO:215 is an extended cDNA sequence for clone RO200 A10 (SEQ ID NO: 157), also referred to as O597S

SEQ ID NO:216 is an extended cDNA sequence for clone RO200 A12 (SEQ ID NO: 158), also referred to as O598S

SEQ ID NO:217 is a full length cDNA sequence for monocarboxylate transporter (MCT3), a sequence having homology with O598S

SEQ ID NO:218 is an extended cDNA sequence for clone RO200 E10 (57881.2; SEQ ID NO: 206), also referred to as O599S

SEQ ID NO:219 is an extended cDNA sequence for clone RO200 G2 (SEQ ID NO: 193), also referred to as O600S

SEQ ID NO:220 is an extended cDNA sequence for clone RO200 B4 (57882.1; SEQ ID NO: 207), also referred to as O601S

SEQ ID NO:221 is a full length cDNA sequence for lysophospholipase I (LYPLA1), a sequence having homology with O601S

SEQ ID NO:222 is an extended cDNA sequence for clone RO201 D1 (57884.2; SEQ ID NO: 208), also referred to as O602S

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for using the compositions, for example in the therapy and diagnosis of cancer, such as ovarian and endometrial cancer. Certain illustrative compositions described herein include ovarian tumor polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (*e.g.*, T cells). An "ovarian tumor protein," as the term is used herein, refers generally to a protein that is expressed in ovarian tumor cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in a normal tissue, as determined using a representative assay provided herein. Certain ovarian tumor proteins are tumor proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with ovarian cancer.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set forth in SEQ ID NO:1-222, antibody compositions capable of binding polypeptides encoded by the polynucleotides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human ovarian cancer.

POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes an ovarian tumor protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term "variants" also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the

sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenes pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS 5*:151-153; Myers, E.W. and Muller W. (1988) *CABIOS 4*:11-17; Robinson, E.D. (1971) *Comb. Theor 11*:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol. 4*:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA 80*:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent

sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as

described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other

polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NO:1-222, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™

technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using any of a variety of well established techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for

tumor-associated expression (*i.e.,* expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to the manufacturer's instructions (and essentially as described by Schena *et al., Proc. Natl. Acad. Sci. USA 93*:10614-10619, 1996 and Heller *et al., Proc. Natl. Acad. Sci. USA 94*:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as ovarian tumor cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.,* an ovarian tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.,* by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook *et al., Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia *et al.*, *Nucl. Acids Res. 16*:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom *et al.*, *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker *et al.*, *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. *et al.* (1995) *Science 269*:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. *et al.* (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid,

or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem. 264*:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to

include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel *et al.* (supra) and Grant *et al.* (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J. 6*:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. *et al.* (1984) *EMBO J. 3*:1671-1680; Broglie, R. *et al.* (1984) *Science 224*:838-843; and Winter, J. *et al.* (1991) *Results Probl. Cell Differ. 17*:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci. 91* :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader

sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci. 81*:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ. 20*:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation. glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction

of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) *Cell 11*:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1990) *Cell 22*:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci. 77*:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol. 150*:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci. 85*:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol. 55*:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. *et al.* (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul. Minn.) and Maddox, D. E. *et al.* (1983; *J. Exp. Med. 158*:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood

by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen. San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. *et al.* (1992, *Prot. Exp. Purif. 3*:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. *et al.* (1993; *DNA Cell Biol. 12*:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc. 85*:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific

mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated

sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR™, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a

sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α-thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the

target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.,* 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter

sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in its entirety), may also be used in the amplification of DNA sequences of the present invention.

BIOLOGICAL FUNCTIONAL EQUIVALENTS

Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

Table 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (–0.4); threonine (–0.7); serine (–0.8); tryptophan (–0.9); tyrosine (–1.3); proline (–1.6); histidine (–3.2); glutamate (–3.5); glutamine (–3.5); aspartate (–3.5); asparagine (–3.5); lysine (–3.9); and arginine (–4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (–0.4); proline (–0.5 ± 1); alanine (–0.5); histidine (–0.5); cysteine (–1.0); methionine (–1.3); valine (–1.5); leucine (–1.8); isoleucine (–1.8); tyrosine (–2.3); phenylalanine (–2.5); tryptophan (–3.4). It is understood that an amino

acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety or well known approaches, several of which are outlined below for the purpose of illustration.

1. ADENOVIRUS

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of

course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5′-tripartite leader (TPL) sequence which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be

generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then

replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9-10^{11} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5′ and 3′ ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into

the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzycska, 1984) is a parovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs (FIG. 2). There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with

all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were

particularly attractive properties for liver-directed gene transfer. Chang *et al.* (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but

it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e. ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present invention.

ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic

antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalactauronase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal $GABA_A$ receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the rat and human sequences) and determination of secondary structure, T_m, binding energy, relative stability, and antisense compositions were selected based upon their

relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris *et al.*, 1997).

RIBOZYMES

Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*,

1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes H-*ras*, c-*fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action

(Woolf *et al.*, 1992). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S. Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells from eukaryotic promoters (*e.g.*, Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992;

Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No. WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5′-end, and phosphoramidites at the 3′-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively

to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-flouro, 2'-o-methyl, 2'-H (for a review see *e.g.*, Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for

eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisziewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used (Christensen *et al.*, 1995).

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, 1995). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this

difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaima *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ulmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremsky *et al.*, 1996; Pardridge *et al.*, 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3′ to 5′ polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The enhanced rate and affinity of hybridization are significant

because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang *et al.*, 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence specificity of PNA recognition. As with DNA hybridization, however, selective recognition can be achieved by balancing oligomer length and incubation temperature. Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a single mismatch within a 16 bp PNA-DNA duplex can reduce the T_m by up to 15°C (Egholm *et al.*, 1993). This high level of discrimination has allowed the development of several PNA-based strategies for the analysis of point mutations (Wang *et al.*, 1996; Carlsson *et al.*, 1996; Thiede *et al.*, 1996; Webb and Hurskainen, 1996; Perry-O'Keefe *et al.*, 1996).

High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an essential RNA template, while the analogous DNA oligomers do not (Norton *et al.*, 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini (Nielsen *et al.*, 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur spontaneously at sequences within chromosomal DNA (Boffa *et al.*, 1995; Boffa *et al.*, 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa *et al.*, 1995) and to

inhibit transcription (Boffa *et al.*, 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene expression. Studies and reviews concerning the use of PNAs as antisense and anti-gene agents include Nielsen *et al.* (1993b), Hanvey *et al.* (1992), and Good and Nielsen (1997). Koppelhus *et al.* (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995), blocking of transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

The present invention, in other aspects, provides polypeptide compositions. Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a contiguous amino acid sequence from an amino acid sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, particularly a polypeptide having the

amino acid sequence encoded by SEQ ID NOs:1-222, or to active fragments, or to variants or biological functional equivalents thereof.

Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NOs:1-222, or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of an ovarian tumor protein or a variant thereof, as described herein. As noted above, an "ovarian tumor protein" is a protein that is expressed by ovarian tumor cells. Proteins that are ovarian tumor proteins also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with ovarian cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of an ovarian tumor protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native ovarian tumor protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ^{125}I-labeled Protein A.

As noted above, a composition may comprise a variant of a native ovarian tumor protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native ovarian tumor protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants

in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See* Merrifield, *J. Am. Chem. Soc.* *85*:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be

selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*, *Gene 40*:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA 83*:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements

responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute *et al. New Engl. J. Med.*, *336*:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium Haemophilus influenza B (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemaglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the LytA gene; *Gene 43*:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology 10*:795-798, 1992). Within a preferred embodiment, a repeat portion of

LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to an ovarian tumor protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to an ovarian tumor protein if it reacts at a detectable level (within, for example, an ELISA) with an ovarian tumor protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as ovarian cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to an ovarian tumor protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the

cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol. 6*:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the

desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y, ^{123}I, ^{125}I, ^{131}I, ^{186}Re, ^{188}Re, ^{211}At, and ^{212}Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers

include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diptheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter *et al.*), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn *et al.*), by

serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler *et al.*).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato *et al.*), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih *et al.*). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison *et al.* discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for an ovarian tumor protein. Such cells may generally be prepared *in*

vitro or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with an ovarian tumor polypeptide, polynucleotide encoding an ovarian tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, an ovarian tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for an ovarian tumor polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen *et al.*, *Cancer Res.* *54*:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with an ovarian tumor polypeptide (100 ng/ml - 100 μg/ml, preferably 200 ng/ml - 25 μg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan *et al.*, Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been

activated in response to an ovarian tumor polypeptide, polynucleotide or polypeptide-expressing APC may be $CD4^+$ and/or $CD8^+$. Ovarian tumor protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, $CD4^+$ or $CD8^+$ T cells that proliferate in response to an ovarian tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to an ovarian tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize an ovarian tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of an ovarian tumor protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup of elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared is such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of

storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human

administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when

administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroetheylene support matrix is described in U. S. Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed

with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In addition, several successful clinical trails examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μm. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is

possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur *et al.* (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is

disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be are easily made, as described (Couvreur *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by reference in its entirety).

VACCINES

In certain preferred embodiments of the present invention, vaccines are provided. The vaccines will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition or vaccine.

Illustrative vaccines may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well

known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems 15*:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA 86*:317-321, 1989; Flexner *et al.*, *Ann. N.Y. Acad. Sci. 569*:86-103, 1989; Flexner *et al.*, *Vaccine 8*:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques 6*:616-627, 1988; Rosenfeld *et al.*, *Science 252*:431-434, 1991; Kolls *et al.*, *Proc. Natl. Acad. Sci. USA 91*:215-219, 1994; Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. USA 90*:11498-11502, 1993; Guzman *et al.*, *Circulation 88*:2838-2848, 1993; and Guzman *et al.*, *Cir. Res. 73*:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer *et al.*, *Science 259*:1745-1749, 1993 and reviewed by Cohen, *Science 259*:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that a vaccine may comprise both a polynucleotide and a polypeptide component. Such vaccines may provide for an enhanced immune response.

It will be apparent that a vaccine may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the vaccine compositions of this invention, the type of carrier will vary

depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.,* polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (*e.g.,* neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.,* glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.,* aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bortadella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA);

aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN-γ, TNFα, IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol. 7*:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science 273*:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO

96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine 14*:1429-1438, 1996) and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature 392*:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med. 50*:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel *et al.*, *Nature Med. 4*:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNFα to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into

dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNFα, CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding an ovarian tumor protein (or portion or other variant thereof) such that the ovarian tumor polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi *et al.*, *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the ovarian tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as ovarian cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host

immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as $CD8^+$ cytotoxic T lymphocytes and $CD4^+$ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow in vivo and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see*, for example, Cheever *et al.*, *Immunological Reviews 157*:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced

into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to an ovarian tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSIS

In general, a cancer may be detected in a patient based on the presence of one or more ovarian tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as ovarian cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, an ovarian tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length ovarian

tumor proteins and portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μg, and preferably about 100 ng to about 1 μg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (*see*, *e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized

on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with ovarian cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter

group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as ovarian cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett *et al.*, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution

containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1μg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use ovarian tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such ovarian tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with an ovarian tumor protein in a biological sample. Within certain methods, a biological sample comprising $CD4^+$ and/or $CD8^+$ T cells isolated from a patient is incubated with an ovarian tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such

as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 μg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of ovarian tumor polypeptide to serve as a control. For $CD4^+$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For $CD8^+$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding an ovarian tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of an ovarian tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the ovarian tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding an ovarian tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding an ovarian tumor protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NOs:1-222. Techniques for both PCR based

assays and hybridization assays are well known in the art (*see*, for example, Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, *51*:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple ovarian tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor

protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to an ovarian tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding an ovarian tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding an ovarian tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding an ovarian tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLE 1

IDENTIFICATION OF CDNAS ENCODING OVARIAN AND ENDOMETRIAL TUMOR PROTEINS

An ovarian/endometrial tumor cell line subtracted library was constructed. A library was prepared from endometrial and ovarian tumor cell lines: EndoTL 391-73 (100% undifferentiated endometrial carcinoma), OTL 298-95 (100% moderately differentiated papillary serous ovarian adenocarcinoma) and OTL 522-24 (30% mesothelial cells/70% poorly differentiated metastatic ovarian adenocarcinoma). This library was subtracted with liver, pancreas, skin, bone marrow, resting PBMC, stomach, and brain cDNA and spiked with eukaryotic elongation factor 1α. Resulting cDNA was cloned into the pcDNA3.1(+) (Invitrogen) vector to generate the ovarian tumor cell line subtraction 4 library (OTCLS4). The OTCLS4 library contained 117,200 clones (background 58,400), with a 1333 bp average insert size (inserts ranged from 200 to 5650 bp).

Thirty clones were sequenced. Of these 12 were full length. The clones may be grouped as follows (SEQ ID NOs are provided in Table 2):

7 Novel

4 Homo sapiens aldehyde dehydrogenase 6 (ALDH6) mRNA

3 Human ferritin heavy chain mRNA, complete cds

2 Human lysyl oxidase gene, partial cds

2 Human mitochondrion, complete genome

1 Homo sapiens aldehyde reductase 1 (low Km aldose reductase) ALDR1) mRNA

1 Homo sapiens chromosome 11q12.2 PAC clone pDJ519o13

1 Homo sapiens chromosome-associated polypeptide C (CAP-C) mRNA

1 Homo sapiens clone 24452 mRNA sequence

1 Homo sapiens dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2) (DPP4 mRNA)

1 Homo sapiens guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1 (GNB2L1), mRNA

1 Homo sapiens heat shock 27kD protein 1 (HSPB1) mRNA

1 Homo sapiens homeo box B2 (HOXB2) mRNA

1 Homo sapiens mRNA for KIAA0865 protein, partial cds

1 Homo sapiens mRNA; cDNA DKFZp564A2416 (from clone DKFZp564A2416)

1 Homo sapiens NADH-ubiquinone oxidoreductase 39kDA subunit mRNA, nuclear gene encoding mitochondrial protein, complete cds

1 Homo sapiens Sk/Dkk-1 protein precursor, mRNA, complete cds

1 Homo sapiens sodium channel, nonvoltage-gated 1 alpha (SCNN1A) mRNA

1 Homo sapiens SRP1 mRNA, partial sequence

1 Homo sapiens zinc finger protein SLUG (SLUG) gene, complete cds

1 Human 28S ribosomal RNA gene

1 Human cofilin mRNA, partial cds

1 Human DNA sequence from clone 967N21 on chromosome 20p12.3-13

1 Human fibroblast collagenase inhibitor mRNA, complete cds

1 Human fibroblast mRNA for aldolase A

1 Human HepG2 3' region MboI cDNA, clone hmd6a06m3

1 Human MAP kinase kinase MEK5c mRNA, complete cds

1 Human mRNA for coupling protein G(s) alpha-subunit (alpha-S1)

1 Human mRNA for KIAA0026 gene, completecds|gi|4808630|gb|AF100620.1| AF100620 Homo sapiens transcription factor-like protein MRGX (MRGX) mRNA, complete cds

1 Human mRNA for KIAA0064 gene, complete cds

1 Human mRNA for KIAA0204 gene, complete cds

1 Human plasminogen activator inhibitor-1 (PAI-1) mRNA, complete cds

1 Human protocadherin 43 mRNA, 3' end of cds for alternative splicing PC43-12

1 Human putative RNA binding protein Koc1 mRNA, complete cds

1 Human TCB gene encoding cytosolic thyroid hormone-binding protein, complete cds

1 Human ubiquitin-homology domain protein PIC1 mRNA, complete cds

Table 2

Ovarian/Endometrial Carcinoma Associated cDNA Sequences

Sequence	SEQ ID NO	Comments
32609	36	*Homo sapiens* aldehyde dehydrogenase 6 (ALDH6) mRNA
32515	4	*Homo sapiens* aldehyde reductase 1 (low Km aldose reductase) (ALDR1) mRNA
32562	29	*Homo sapiens* Chromosome 11q12.2 PAC clone pDJ519o13
32523	9	*Homo sapiens* chromosome-associated polypeptide C (CAP-C) mRNA
32551	24	*Homo sapiens* clone 24452 mRNA sequence
32518	6	*Homo sapiens* dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2) (DPP4) mRNA
32534	13	*Homo sapiens* guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1 (GNB2L1), mRNA
32507	2	*Homo sapiens* heat shock 27kD protein 1 (HSPB1) mRNA
32533	12	*Homo sapiens* homeo box B2 (HOXB2) mRNA
32565	20	*Homo sapiens* mRNA for KIAA0865 protein, partial cds
32553	19	*Homo sapiens* mRNA; cDNA DKFZp564A2416 (from clone DKFZp564A2416)
32561	28	*Homo sapiens* NADH-ubiquinone oxidoreductase 39kDa subunit mRNA, nuclear gene encoding mitochondrial protein, complete cds
32510	3	*Homo sapiens* Sk/Dkk-1 protein precursor, mRNA, complete cds
32546	16	*Homo sapiens* sodium channel, nonvoltage-gated 1 alpha (SCNN1A) mRNA
32559	27	*Homo sapiens* SRP1 mRNA, partial sequence
32506	1	*Homo sapiens* zinc finger protein SLUG gene, complete cds
32519	7	Human 28S ribosomal RNA gene
32602	22	Human cofilin mRNA, partial cds
32569	31	Human DNA sequence from clone 967N21 on chromosome 20p12.3-13
32525	10	Human ferritin heavy chain mRNA, complete cds
32557	26	Human fibroblast collagenase inhibitor mRNA, complete cds
32517	5	Human fibroblast mRNA for aldolase A
32568	30	Human HepG2 3' region MboI cDNA, clone hmd6a06m3
32548	17	Human lysyl oxidase gene, partial cds
32520	8	Human mitochondrion, complete genome
32617	23	Human mRNA for coupling protein G(s) alpha-subunit (alpha-S1)

32572	32	Human mRNA for KIAA0026 gene, complete cds|gi|4808630|gb|AF100620.1|AF100620 *Homo sapiens* transcription factor-like protein MRGX (MRGX) mRNA, complete cds
32600	21	Human mRNA for KIAA0064 gene, complete cds
32537	14	Human mRNA for KIAA0204 gene, complete cds
32552	25	Human plasminogen activator inhibitor-1 (PAI-1) mRNA, complete cds
32615	39	Human protocadherin 43 mRNA, 3' end of cds for alternative splicing PC43-12
32613	38	Human putative RNA binding protein Koc1 mRNA, complete cds
32610	37	Human TCB gene encoding cytosolic thyroid hormone-binding protein, complete cds
32539	15	Human ubiquitin-homology domain protein PIC1 mRNA, complete cds
32619	40	Novel
32576	33	Novel
32608	35	Novel
32607	34	Novel
32620	41	Novel
32550	18	Novel
32529	11	Novel

Using the methods outlined above, an additional 162 clones were isolated and sequenced. The cDNA sequences are shown in SEQ ID NO:42-203.

SEQ ID NO:204-209 represent additional clones from the OTCL S4 library. SEQ ID NO:206 (clone 57881), 208 (clone 57884), 107 (clone R0199:A07) and 80 (clone R0198:F02) represent novel sequences. The remaining sequences are shown in Table 3, which includes additional results from homology searches.

Table 3

Sequence	SEQ ID NO	Comments
57877	204	H. Sapiens novel gene from PAC 117P20, chromosome 1
57879	205	Urokinase plasmingen activator surface receptor (uPAR)
57882	207	Lysophospholipase 1 (LYPA1)
57888	209	IGF-II mRNA binding protein 3 (IMP-3) mRNA
R0198:H03	99	*Homo sapiens* laminin
R0199:B03	111	Human cyclin protein gene, complete cds
R0200:A12	158	*Homo sapiens* monocarboxylate transporter (MCT3) mRNA
R0199:C12	125	Unigene: Hs93379
R0200:A10	157	Human mRNA for KIAA0101 gene, complete cds
R0198:D01	61	Unigene: Hs42116
R0200:C02	164	Human proliferating cell nuclear antigen (PCNA) gene
R0200:G02	193	*Homo sapiens* Xq28 BAC RP5–1014016

EXAMPLE 2

ANALYSIS OF cDNA EXPRESSION USING MICROARRAY TECHNOLOGY

In additional studies, sequences disclosed herein were found to be overexpressed in specific tumor tissues as determined by microarray analysis. Using this approach, cDNA sequences are PCR amplified and their mRNA expression profiles in tumor and normal tissues are examined using cDNA microarray technology essentially as described (Shena *et al.*, 1995). In brief, the clones are arrayed onto glass slides as multiple replicas, with each location corresponding to a unique cDNA clone (as many as 5500 clones can be arrayed on a single slide, or chip). Each chip is hybridized with a pair of cDNA probes that are fluorescence-labeled with Cy3 and Cy5, respectively. Typically, 1μg of polyA$^+$ RNA is used to generate each cDNA probe. After hybridization, the chips are scanned and the fluorescence intensity recorded for both Cy3 and Cy5 channels. There are multiple built-in quality control steps. First, the probe quality is monitored using a panel of ubiquitously expressed genes. Secondly, the control plate also can include yeast DNA fragments of which complementary RNA may be spiked into the probe synthesis for measuring the quality of the probe and the sensitivity of the analysis. Currently, the technology offers a sensitivity of 1 in 100,000

copies of mRNA. Finally, the reproducibility of this technology can be ensured by including duplicated control cDNA elements at different locations.

A total of 428 clones from the OCTLS4 library were analyzed on Ovarian Chip-3. The following table, Table 4, provides a list of probes used to interrogate these clones. A total of 16 clones were identified which showed at least 2-fold overexpression in ovarian tumors when compared to non-ovarian essential normal tissues and had a mean non-ovarian essential normal tissue expression of less than 0.2. These clones are represented by SEQ ID NO:204-209 and by SEQ ID NO:61, 99, 111, 125, 157, 158, 164 and 193.

Table 4

Tissue	Clone ID	Microarray ID	Tumor information
Ovarian tumor Adrenal gland normal	261A SPACT37	391cy3 391cy5	Stage IIIC
Ovary tumor Skin normal	264A 396A	392cy3 392cy5	Stage IIIC
Ovary tumor Thymus normal	265A SPACT56	393cy3 393cy5	Stage IIIC
Ovary tumor Bronchus normal	288A 600C	394cy3 394cy5	Stage IIIC
Ovary tumor	854A 785B	395cy3 395cy5	
Ovary tumor Bone normal	855A 407B	396cy3 396cy5	Grade III, Stage IA
Ovary tumor Peritoneum epithelium normal	856A 484A	397cy3 397cy5	Serous papillary
Ovary tumor Pituitary gland	603A SPACT52	398cy3 398cy5	Metastatic adenocarcinoma, Grade III, Stage III
Ovary tumor Skeletal muscle normal	857A SPACT40	399cy3 399cy5	Papillary serous cystadenocarcinoma Grade III, Stage IB
Ovary tumor Stomach normal	385A SPACT55	400cy3 400cy5	Papillary serous adenocarcinoma
Ovary tumor Spleen normal	392A SPACT54	401cy3 401cy5	Papillary serous neoplasm, Stage 1C
Ovary tumor Pancreas normal	858A 862A	402cy3 402cy5	Papillary serous cystadenocarcinoma Grade II-III, Stage IA

Ovary tumor Ovary normal	859A S27	403cy3 403cy5	Papillary serous adenocarcinoma Grade II-III, Stage IIB
Ovary tumor Spinal cord normal	605A SPACT45	404cy3 404cy5	Serous borderline tumor, stage IIIC
Ovary tumor Heart normal	495A SPAAm1	405cy3 405cy5	Papillary serous carcinoma, Grade II, Stage III
Ovary tumor Ovary normal	381C S7	414cy3 414cy5	Mucinous adenocarcinoma, Grade I, Sage IB
Ovary tumor Ovary normal	382A S449A	416cy3 416cy5	Mucinous adenocarcinoma
Ovary tumor metastases Small intestine normal	428B SPACT53	417cy3 417cy5	Mucinous adenocarcinoma
Ovary tumor Esophagus normal	491A 502B	418cy3 418cy5	Endometriod adenocarcinoma
Ovary tumor Colon normal	335A 199A	419cy3 419cy5	Endometriod adenocarcinoma Grade II, Stage II
Ovary tumor Thyroid gland normal	494A SPACT46	421cy3 421cy5	Adenocarcinoma Grade III, Stage II-III
Ovary tumor PBMC (resting)	860A 783A	42cy3 422cy5	Endometriod adenocarcinoma Grade II-III, Stage IIIC
Ovary tumor Aorta normal	604A 415A	423cy3 423cy5	Clear cell carcinoma
Ovary tumor Trachea normal	607A 776A	424cy3 424cy5	Clear cell, Stage IA
Ovary tumor Trachea normal	S25 CT25	425cy3 425cy5	Granulosa cell tumor, Stage IA
Ovary tumor Pancreas normal pool	S22 PAN2000	426cy3 426cy5	Granulosa cell tumor, Stage IA
Ovary tumor Breast (HMEC) normal	386A S92	427cy3 427cy5	Germ cell tumor, Stage I
Ovary tumor Bladder normal	602A 328B/C	429cy3 429cy5	Papillary serous carcinoma, Grade III, Stage IIIB
Ovary tumor Bone marrow normal	S23 SPACT49	430cy3 430cy5	Papillary serous adenocarcinoma Grade III, Stage IIIC

Ovary tumor Lung normal	606A SPAAm2	428cy3 428cy5	Papillary serous cystadenocarcinoma Grade II, Stage IIIB
Ovary tumor metastases Kidney normal	383A 302B	431cy3 431cy5	Metastatic papillary adenocarcinoma, Grade III, Stage IIIA
Ovary tumor metastases PBMC (activated)	384A S40.782A	423cy3 423cy5	Papillary serous adenocarcinoma Grade II, Stage IIIB
Ovary tumor metastases Ovary tumor match with CY3	426A 603A	433cy3 433cy5	Papillary serous adenocarcinoma Grade III, Stage IIIB
Ovary tumor metastases Liver normal	429A 270B	434cy3 434cy5	Papillary adenocarcinoma Grade III, Stage III
Ovary tumor Brain normal	427A SPACT50	435cy3 435cy5	Papillary serous adenocarcinoma Grade III, Stage IIIC
Ovary tumor Bone normal	855A 407B	436cy3 436cy5	Grade III, Stage IA
Ovary tumor Spinal cord normal	605A SPACT45	437cy3 437cy5	Serous borderline tumor, Stage IIIC
Ovary tumor Heart normal	495A SPAAm1	438cy3 438cy5	Papillary serous carcinoma, GradeII, Stage III
Ovary tumor Ovary normal	381C S7	439cy3 439cy5	Mucnous adenocarcinoma, Grade I, Stage IB

EXAMPLE 3

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed:

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NO: 1-222;

(b) complements of the sequences provided in SEQ ID NO: 1-222;

(c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NO: 1-222;

(d) sequences that hybridize to a sequence provided in SEQ ID NO: 1-222, under moderately stringent conditions;

(e) sequences having at least 75% identity to a sequence of SEQ ID NO: 1-222;

(f) sequences having at least 90% identity to a sequence of SEQ ID NO: 1-222; and

(g) degenerate variants of a sequence provided in SEQ ID NO: 1-222.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) sequences encoded by a polynucleotide of claim 1; and

(b) sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and

(c) sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

3. An expression vector comprising a polynucleotide of claim 1 operably linked to an expression control sequence.

4. A host cell transformed or transfected with an expression vector according to claim 3.

5. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide of claim 2.

6. A method for detecting the presence of a cancer in a patient, comprising the steps of:

(a) obtaining a biological sample from the patient;

(b) contacting the biological sample with a binding agent that binds to a polypeptide of claim 2;

(c) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(d) comparing the amount of polypeptide to a predetermined cut-off value and therefrom determining the presence of a cancer in the patient.

7. A fusion protein comprising at least one polypeptide according to claim 2.

8. An oligonucleotide that hybridizes to a sequence recited in SEQ ID NO: 1-222 under moderately stringent conditions.

9. A method for stimulating and/or expanding T cells specific for a tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

(a) polypeptides according to claim 2;

(b) polynucleotides according to claim 1; and

(c) antigen-presenting cells that express a polypeptide according to claim 2,

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

10. An isolated T cell population, comprising T cells prepared according to the method of claim 9.

11. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

(a) polypeptides according to claim 2;

(b) polynucleotides according to claim 1;

(c) antibodies according to claim 5;

(d) fusion proteins according to claim 7;

(e) T cell populations according to claim 10; and

(f) antigen presenting cells that express a polypeptide according to claim 2.

12. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 11.

13. A method for the treatment of a cancer in a patient, comprising administering to the patient a composition of claim 11.

14. A method for determining the presence of a cancer in a patient, comprising the steps of:

(a) obtaining a biological sample from the patient;

(b) contacting the biological sample with an oligonucleotide according to claim 8;

(c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(d) compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

15. A diagnostic kit comprising at least one oligonucleotide according to claim 8.

16. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

17. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of: (i) polypeptides according to claim 2; (ii) polynucleotides according to claim 1; and (iii) antigen presenting cells that express a polypeptide of claim 2, such that T cell proliferate;

(b) administering to the patient an effective amount of the proliferated T cells,

and thereby inhibiting the development of a cancer in the patient.

SEQUENCE LISTING

<110> Corixa Corporation
Xu, Jiangchun
Pyle, Ruth
Secrist, Heather

<120> COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER

<130> 210121.501PC

<140> PCT
<141> 2001-03-20

<160> 222

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 595
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(595)
<223> n = A,T,C or G

<400> 1
ccttttttttt ttttttaaa tcaaaactgt ttattgtaaa aaaaacttga aaattgtttt 60
ttaaaaaaga aacattgatt tcacaagtct tcaggttgtt tatagacata gctatagaca 120
acatctcagt ttcatacaga actcattcaa tcatataaaa ataaacacaa atttacattg 180
actcatcaac tatacaattt aaaaaggcac ttggaagggg tattgtatta ttgcatttgt 240
ggtatgcatt tgaaatagtt taagtacatt aatgaatttg taagaatcct cttttgcact 300
tattcccatc tttaattaat tttcaaaaat tattaaaatg ttttaaaata gtaagacaat 360
ggagcatgcg ccaggaatgt ttcaaagcta atctttccct cctcccccaa ggcacatact 420
gttaattggg caaaaacaaa aacaaacaaa aatactttta atacattctc ctgggggttg 480
gnncttggna attttttttt ccccttttaaa aatatacctt taangcnctc aggtaatcaa 540
aaaaaaggct ttagtcacaa ntggcnaccc gnccaaccca ctngcacngg nntan 595

<210> 2
<211> 1700
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1700)
<223> n = A,T,C or G

<400> 2
aaaagcgcag ccgagcccag cgccccgcac ttttctgagc agacgtccag agcagagtca 60
gccagcatga ccgagcgccg cgtccccttc tcgctcctgc ggggccccag ctgggacccc 120
ttccgcgact ggtacccgca tagccgcctc ttcgaccagg ccttcgggct gccccggctg 180

ccggaggagt ggtcgcagtg gttaggcggc agcagctggc caggctacgt gcgccccctg 240
ccccccgccg ccatcgagag ccccgcagtg gccgcgcccg cctacagccg cgcgctcagc 300
cggcaactca gcagcggggt ctcggagatc cggcacactg cggaccgctg gcgcgtgtcc 360
ctggatgtca accacttcgc cccggacgag ctgacggtca agaccaagga tggcgtggtg 420
gagatcaccg gcaagcacga ngagcggcag gacgagcatg gctacatctc ccggtgcttc 480
acgcggaaat acacgctgcc ccccggtgtg gaccccaccc aagtttcttc tccctgtccc 540
ctgagggcac actgaccgng gaggncccca tgcccaagct agccacgcag tccaacgaga 600
tcaccatncc agtnaccttc nantngcggg cccagcttgg gggnccanaa nctnnnaaaa 660
tccnataaga ntggccgcca anaaannnct tannnccggg atgcccaccc cttgntgcng 720
ccnntgggtn gggccttccc ccnccnccng gggggnnntt tnnananann nanntnnggn 780
nnnnnnnnaa aaggnnnnna ngnnnccccn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngngg ngnnnnnnnn 900
nnnnnntnnn nnnnnnnccn cnngnnnnnn nnngnnnnnn nnnnnnnnnn nnnnnnnnnn 960
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngggnntnn 1020
tnntnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnngn nncncnnnnn nnnnnnnnnn 1080
nnnnnnggnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140
nnnnngnngn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1440
nnnnnnncncn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1500
nnnnnnnccc ccccccccccc ccccccccccc ccccccccccc cnnnnnnnnn nnnnnccccc 1560
ccccccccccc nnttttttnc ccccccccccc ccccccccnn nnnnnnnnnn nnnnnnnnnn 1620
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1680
nnnnnnnnnn nnnnnnnngn 1700

<210> 3
<211> 583
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(583)
<223> n = A,T,C or G

<400> 3
ccttttttttt ttttttttga tattaaatgt taaattttat ttcaaaaact atcacagcct 60
aaagggaaat ataatttaag cattaagata gtacatttca gaaaataagc tagtattttc 120
atgttacatt ttaggtacct atcatttgtc attccaagag atccttgcgt tctagactct 180
anaattaaat ggggtaaagg gttatgcttt taagaactat aagctgaaat gatttacttc 240
agttcaatat agaatatttg tcagtcaaga taacaatcaa tgtgtcaaaa atttacataa 300
caagaggaaa aataggcagt gcagcacctt tagaaaaata attaaaagtt tcattgcatt 360
tacangnaag tgccacactg agaatttaca atacagtaat ttactgcaat cacaggggag 420
ttccataaag aaacaaagct cttcactcca ggtttttgga angggtatt ggaagcttaa 480
ctgaaacccc aaaacntggt tantcctnng aatgagttga tgaaaggcat aaaaagggtt 540
cttagccctn ttntntaaaa gggggccccg ctttgggaaa cng 583

<210> 4
<211> 448
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(448)
<223> n = A,T,C or G

<400> 4
cctttttttt ttttttttca caaaagcact ttttatttga ggcaaagaga agtcttgctg 60
aaaggattcc agttccaagc agtcaaaact caaccgttag tggcactatt ttgacctggt 120
agattttgct tctctttggt canaaaaggg tattcaggtt gtactttccc cagcagggta 180
gaaagaaggg caaagcaaac tggaagagac ttctactcta ctgacagggc tcttgagatc 240
caacatcaag ctagacacgc cctcgctggc cactctacag gttgctgtcc cactgctgag 300
tgacacaggc catactacat ttgcaaggaa aaaaatgagg caagaaacac aggtataggt 360
cacttgggga cgagcaggca accacagctt caaaactctt catggaaggg gtaatccttg 420
nggggaggna cagctcaagt cgaccggc 448

<210> 5
<211> 2067
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(2067)
<223> n = A,T,C or G

<400> 5
ccgaggctaa atcggctgcg ttcctctcgg aacgcgccgc anaggggtc ctggtgacga 60
gtcccgcgtt ctctccttga atccactcgc cagcccgccg ccctctgccg ccgcaccctg 120
cacacccgcc cctctcctgt gccaggaact tgctactacc agcaccatgc cctaccaata 180
tccagcactg accccggagc agaagaanga gctgtctgac atcgctcacc gcatcgtgga 240
cctggcaagg gcatcctgnc tgcagatgag tccactggga gcattgncaa gcggctgcat 300
tccattggca ccgagaacac cgaggagaac cggcgcttct accgccagct gctgctgaca 360
gctgacgacc gcgtgaaccc ctgcattggg ggtgtcatnc tcttccatga gacactctnc 420
cagaaggcgg atgatgggcg tcccttcccc caagttatca aatccaaggg cggtgttggg 480
gggcatcaag gtagacaagg gcgnggtccc cctggcaggg gacaaatggn gagactacca 540
cccaaagggt tggatgggct gtctgaancn ctgngcccag nacnaanaan gacggagctg 600
acttccccaa ntggngtttg ngtgctnaaa aattggggaa aacaaccccc ctnaaaccct 660
tcngcattna tggaaaaatn cccaatgttn tgggnccntn angccnngnt ntnccannnn 720
naangggatt tnngcnccnt nnnnggancc nnnnanancc ncccccttgng gggggnaaca 780
tnnaannttn naanngnnnn gnncnnnnnn ngnnnnancn nnannanaan ggnnnnnnng 840
nnnntgnnnn nnnncnnann anggnncnnn nnnnnngngn gancgcnnnc cnnnnnnnng 900
nnnancnngn naaagngana ccnngnatnn tnnnnangnn ncnanannnn gngtnnnnnn 960
nnannnnnnn nnnnngnggg gcgnnngcng nnnccnnngn ngnnngnnnn nnnnnnnnnc 1020
nggnnaaaaa nnnnccnccc cnncnnnnnn cncnnnnnna annnnntnnn nnnnncnnnc 1080
ccnngnannc nnngnnnnnn gnannnnnnn gngnacgnnn nnngnnnngn ngnnncnnnn 1140
ntnnnnnncg nnnnnnngnn nnannnnnnn nnnnanannn nnannnnnnn agngngnnng 1200
nggggnngnt nntngnatgn ncnnnnnnnn nnnnnnnncn nnntnntnnn nnnnnnnnnn 1260
nnnnnannng nnnnngnncn nnnangnnnn nnnnnnnnng nnnnnnnnnn nnngnnnnnn 1320
nnnnnnnnnn gnnnnnnncn cgnnnnnnnn nggngnaaaa aaaatnncgn nctnnnnnng 1380
ngngnnnnnn nnnnangnga aanannnnnn nnnnnnnnnn nnnnnnnnnn nnnnngnagn 1440
nanannnnnn gnngnnnnnn nnngnnnnnn nnnnnngnnn nnnncnncgn ngnncnnncn 1500
nnnnnnnnnn nncncaannn nnnncncncc nannnnnnnn nnnannnncg ncngnnaann 1560
nnannnannn annccnnnnc nnannnncnn nnannnngnn nnnngnnngn nnnnnnannn 1620
nnnnnnnnnn ncnncgnnng ganngnnnnn nnnnnnnnnn nnnntnnnna nggnggnnnn 1680
nngnnggnan nnnngnnnnn nnncnnaann nnnngnnnng cgngngnnnn nnggngnnnn 1740
nnncnnncnn nnnnngnnnn annnnnnnnn nnnnnnnnnn nnnnnnnntn nngntnnnnn 1800
nnnnnnnnnn nncnnnncnn nnnnnnnngn nnnnnnngnn nnnnnnngnn gngnnancng 1860
tngannanan aanngncgaa naagtnngng nnnnnnngnn gnngnncnnc ccnanncnna 1920
ntnnncgnan nngntgagan nnangnggnn aantcnnngg ccnncngncn ngnnngnnca 1980
nnacncggnn ngnnncnggn nngaananan ggggggannn nnnncngggg nccncnnnnn 2040
nnnnannana ngaaaaanaa anagcgn 2067

<210> 6
<211> 643
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(643)
<223> n = A,T,C or G

<400> 6
ccttttttttt ttttttttttt tctgaaaaaa tgaaggcaca tttattaaat gactgggaga 60
aattccatag tatgtagaat gggaataata atacataaca ttgtatttta tgttccattt 120
tttaaaatga gtccaaggaa gttaaaatat tcttttaatt aagacactca aagaaatgaa 180
ataagaaaaa ttgatgcaag gactccttca agttaanatt tgtgatacaa atattttcat 240
cttttaacag ggcaagctga tgtgttcaca tctcagtttc aagctgcctc tttcactagg 300
aacatcagta tttttttttta aaagcacatt tacaatgctt tcccatcacc cttgctgtgt 360
ttttgtagca cctatagcca taactggcac ctgggggcct gcgttgctgg cantttccct 420
tacatttctt tggagtcttt tcaactgctg ggggtttact taaaagtcag tgctttgcat 480
atttgatttc ctganantgn ttgaatagnn tttttaaaaa aatgngcagg ctgggtggga 540
cannttttttt ncaagggaat ganannancn tgctnnggtt ggntngcttg gaatgggtcc 600
aaccnnnnct nnttttnttc ccnancccctt nccngcccng cct 643

<210> 7
<211> 123
<212> DNA
<213> Homo sapien

<400> 7
cctcgcccgt cacgcaccgc acgttcgtgg ggaacctggc gctaaaccat tcgtagacga 60
cctgcttctg ggtcggggtt tcgtacgtag cagagcagct ccctcgctgc gatctattga 120
aag 123

<210> 8
<211> 655
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(655)
<223> n = A,T,C or G

<400> 8
gtaaaaccca gcccatgacc cctaacaggg gccctctcag ccctcctaat gacctccggc 60
ctagccatgt gatttcactt ccactccata acgctcctca tactaggcct actaaccaac 120
acactaacca tataccaatg atgggcgcga tgtaacacga gaaagcacat accaaggcca 180
ccacacacca cctgtccaaa aaggccttcg atacgggata atcctattta ttacctcaga 240
agttttttc ttcgcaggat ttttctgagc cttttaccac tccagcctag ccctaccccc 300
ccaactagga gggcactggc ccccaacagg catcaccccg ctaaatcccc tagaagtccc 360
cgtnctaaac acatncgtat tactggnatg aggagtatca atcacctgag ctcaccatag 420
tctaatagaa aacaaccgaa accaaataat tcaagcactg cttattacaa ttttactggg 480
tctctatttt acccttctac angcctcana atactttcga gtcttcctta acatttccga 540
cggcatctac cggttaacat tttttgtagc cacaaggttt cacggacntt ccctatcatt 600
ggctnacttt tcttactatt ggttattcgc caataaaatt cacttttnnt ccnag 655

<210> 9
<211> 663

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(663)
<223> n = A,T,C or G

<400> 9
ccggagccga aacaccggta ggagcgggga ggtgggtact acacaaccgt ctccagcctt 60
ggtctgagtg gactgtcctg cagcgaccat gccccgtaaa ggcacccagc cctccactgc 120
ccggcgcaga gaggaagggc cgccgccgcc gtcccctgac ggcgccagca gcgacgcgga 180
gcctgagccg ccgtccggcc gcacggagag cccagccacc gccgcagaga ctgcaagtga 240
ggaacttgat aatagaagtt tagaagagat tttgaacagc attcctcctc ccccgcctcc 300
agcaatgacc aatgaagctg gagctcctcg gcttatgata actcatattg taaaccagaa 360
cttcaaatcc tatgctgggg agaaaattct gggacctttc cataagcgct tttcctgtat 420
tatcgggcca aatggcagtg gcaaatccaa tgttattgat tctatgcttt ttgtgtttgg 480
ctatcgagca caaaaaataa gatctaaaaa actctcagta ttaatacata attcttgatg 540
aacnccaagg acnttcagaa ttgnacagta naaagttctt tttcaaaaaa taattggtta 600
agggaagggg tngattttga aancntttct taacnnaant ttttngnttt cccaaacggc 660
tnt 663

<210> 10
<211> 654
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(654)
<223> n = A,T,C or G

<400> 10
gtcggggttt cctgcttcaa cagtgcttgg acggaacccg gcgctcgttc cccaccccgg 60
ccggccgccc atagccagcc ctccgtcacc tcttcaccgc accctcggac tgccccaagg 120
cccccgccgc cgctccagcg gccgcgcagc caccgccgcc gccgccgcct ctccttagtc 180
gccgccatga cgaccgcgtc cacctcgcag gtgcgccaga actaccacca ggactcagag 240
gccgccatca accgccagat caacctggag ctctacgcct cctacgttta cctgtccatg 300
tcttactact ttgaccgcga tgatgtggct ttgaagaact ttgccaaata ctttcttcac 360
caatctcatg aggagaggga acatgctgag aaactgatga agctgcagaa ccaacgaggt 420
ggccgaatct tccttcagga tatcaagaaa ccagactgtg atgactggga gagngggntg 480
aatgccnngg agggggcatt acatttggaa aaaaatgtga atcaagcact actggaactg 540
caccaactgg ccctgacaaa atgacccccа tttgngtgac tttnttgaaa ccatttactt 600
gatgagcagg ggaaancctt cnnnaatggg gngacacgng accaacttgc gnnt 654

<210> 11
<211> 653
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(653)
<223> n = A,T,C or G

<400> 11
ttttttttt ttttttttt tatgggaaac tgctctttat ttagaccttt gggacaaaat 60
taactttggt cacatattac ttaaaaaaaa atccagtttt acatatttct aaatagatag 120

aactaaatga tcagagaatt tcttctgtaa aaattggcca aattttatca aaaatctaac 180
atacgataca atccaaatta taaaaagact acttgggatc ataatattcc aaatgtatga 240
cagttataac tccatcttaa caagtgtgaa aagtacttgc tctcatgttg ctttggtcca 300
aaagagtaga gctaactcag taacaggaaa ctaagtaccc aatcttttgc caaaattaat 360
ttagattgtg actggcagca naaatatcca taatgaacag ctctactata acaaagaata 420
attaaagaat acttttcgtg aacatatcac aggtcaaata cattttata agagaaaaat 480
atgaaggaaa tgatnaaata gctntcncaa acaaaaagga agcatttncc ccntaagggg 540
aattaanagg gtggatgatg cttatatgaa angaagtnga annengnttt atttcttatt 600
tttccactct tanctttcaa aatnggtttg ncatgcccta aagngaancc ngg 653

<210> 12
<211> 375
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(375)
<223> n = A,T,C or G

<400> 12
ttttttttt ttttttttt ttttttggna ttataaanac atttatttaa tctatgaaaa 60
taatgnacaa taaatacttt ccccttttcc tattattaaa naattttaat aaataatnta 120
cagtctaaaa cataaaaaag aggaaaatag gnccctctag ttatttttaa naaagncccc 180
ctanagttta attattcctg anatttcatt ggaaggagtc taccaaacgg aatttttctg 240
ngngaatttt aaaanataac cgagtgccca atattttaga agaagaagaa aggaagngga 300
ttaaacgcta attcagtaat acctgaattt tagcaaaaca cataagtcta tgcgactgag 360
ggngggagan gntcg 375

<210> 13
<211> 658
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(658)
<223> n = A,T,C or G

<400> 13
ctctctcttt cactgcaagg cggcggcagg agaggttgtg gtgctagttt ctctaagcca 60
tccagtgcca tcctcgtcgc tgcagcgaca cacgctctcg ccgccgccat gactgagcag 120
atgacccttc gtggcaccct caagggccac aacggctggg taacccagat cgctactacc 180
ccgcagttcc cggacatgat cctctccgcc tctcgagata agaccatcat catgtggaaa 240
ctgaccaggg atgagaccaa ctatggaatt ccacagcgtg ctctgcgggg tcactcccac 300
tttgttagtg atgtggttat ctcctcagat ggccagtttg ccctctcang ctcctgggat 360
ggaaccctgc gcctctggga tctcacaacg ggcaccacca cgaggcgatt tgtgggccat 420
accaaggatg tgcttgagtg tggccttctc tttgacaacc cggcagattg nctttggat 480
ctcnanaata aaaccatcaa nctattgaat accctgggng tggtgcaaat cccntgtcca 540
ngaaganaac cncttcanaa ngggggtctt tgtgnnccnt ttttnnccca acncaacaac 600
cctnttattn nntncctngg gttggaaaan ctggcnnggn tnganccggn tnactggg 658

<210> 14
<211> 686
<212> DNA
<213> Homo sapien

<220>

<221> misc_feature
<222> (1)...(686)
<223> n = A,T,C or G

<400> 14
ccttttttttt ttttttttt ttttttttt aacattatac tgncattttt atcataacaa 60
tataaacaat ttttatcatc atcctgaata ttactttata aanatatata ttttaaaagg 120
ntttcaaaac atttttcaac ccagcatttg agaataaagc attaagagtt ttgnatacag 180
taacacattc atgngataag ngnatgaatt tacaaccata cataatatgg atatatggat 240
atatatttat ataaaaaaca aacttggcca naagttaagg ntacctacna agttgtccaa 300
gtaaattatg cttggcaaaa caattataaa attcaaatca cacatgcatt tttaaatcat 360
ctaaatcact gcaaacaang gtcaagcatt ccaaangttt taaaatnang ggggangang 420
ggaancnggc cctccaannt taaagggccc gtttaaaacc cccttgaccc ccccccccaca 480
ggngnttttt aactnccncc catttntgtt gtttgnncnt ttcnccgggg ccttctttgg 540
cccttggang gggccncccc cccctgggcc ttccnaaata aaagggagga aaannngnntt 600
cccacgnccc cccccgnatg natnctctcc tntaaaaaaa ngggngggnc gnganncтaa 660
nnggagnggt ttggcnaanc acttct 686

<210> 15
<211> 725
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(725)
<223> n = A,T,C or G

<400> 15
ccttttttttt ttttttttgat ttttacaaat attgnttatt ttaatgaagc tggtacagac 60
aatgtccatt taaaacccat atcccaggcc aaaaagtaca aataaaatca aaaagagcag 120
tgttctgntg tattcatttc tgnatgtata gctttattaa ttngctaatg aaaattanaa 180
cttttctggg atcttctgac aagattttta aaaaatctta aaatgccttt tcttcagtga 240
aggcactttt ggagttncca ataaaggggn ccccccctnc catcttnact tnaacctgat 300
attnntnttg tgnnggggggg ggngggngaa attttaaaaa tatnttaatt taaggaaagg 360
ncatttttтc acagtctaag ttctntgnaa aacttncatt ttcccacnga aagnganagt 420
tnangaannc ccccnngggc ncnccccacc ntgnggggca anttgnaaan tnattatnga 480
acncttggta ttgnttgaat tntttntgnt aacgnnnaat tgcgtgnaag aangctatcg 540
ttnctgtaaa aaaaagggga aacttttnct atantntccn ntannttctt tttanaaacc 600
ccnacccccc ctaaatgtga nccnccgatn ttttnccggg gntggatntt nntcngccct 660
tcncncnccg ccctttttтt anacgccnat ttatattttn taantttatn taanttctca 720
tntct 725

<210> 16
<211> 196
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(196)
<223> n = A,T,C or G

<400> 16
cngaaggtng cctncaccct ggcatcctcc cctccttccn acttntgccc ccaccccatg 60
tctctgtcct tgtcccagcc aggccctgct ccctctccag ccttgacagc ccctcccccт 120
gcctatgcca ccctgggccc ccgcccatct ccaggggget ctgcaggggc cagttcctcc 180
gcctgtcctc tggggg 196

<210> 17
<211> 667
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(667)
<223> n = A,T,C or G

<400> 17
cagccgtgaa actggaaagt cattttgatg actgatgtga tacatccaga ggtaaaatgc 60
atttaaacat attaaagtat ttgccaaaga tacaattttc ttgctgacat aaaaatcaca 120
caaacaagtc ccccccaaac cacaactgtc tctcaaatag cttaaaaaaa ttgaaaaaca 180
ttttaggatt tttcaagttt tctagatttt aaaaagatgt tcagctatta gaggaatgtt 240
aaaaatttta tattatctag aacacaggaa catcatcctg ggttattcag gaatcagtca 300
cacatgtgtg tgtgtctgag atatagtcta aattagcaaa gcacatagta ttacatactt 360
gaggggttgg tgaacaaagg aaaaatatac tttctgcaaa accaangact gtgctgcgta 420
atgagacagc tgtgatttca tttgaaactg tgaaaccatg tgccataata gaattttgag 480
aattttgctt ttacctaaat tcaagaaaat gaaattacac ttttnagtta gnggnggctt 540
aacataattt tttctatntt aacccgtatt naaatctcaa gtaagaattt nccgtggccc 600
gaaacttgtt angggggaat tttaaaaggg cctcgcattc cgggttacat ggcntanaan 660
tggaagg 667

<210> 18
<211> 1493
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1493)
<223> n = A,T,C or G

<400> 18
ccccatttct ccattttgtg gaccaagcca tcctgagggc atggacattg tctctgagga 60
aattggggcc acccttaaga taccaagaaa agctcctgcc catggtccca ctggaaatgg 120
actctgctga gcaaagccac cagttgaaga gaacagaatc cacacctgca ttgaatacct 180
gtttctccat gtgtatcgtc tctgagatta ccttcttgcc ctttccaaca ccttagtgat 240
tcctcaattt ctcccccatt gggaaggcca tagggcatta actgaaggaa ctgacctctc 300
tccttttcct gtaccttaa cctttagtct gtcaaggaaa acccttagga cctctgaatc 360
aagaggactg agtttgtggg tgaaccttga aggtgctctt tctgctacaa gggccctggg 420
agatagcatg ggacgtgcat tgagaagcca gcctcagacc ttagcttgaa gcancttgag 480
gccagaccta ctgtacctca gcatcttgct aggaggcatg gaagtgatct atcctgccag 540
gaggcctcag agtgatctgt cctgccagga ggggtgagag tgatctgtcc tgtgaggcat 600
ttaggggctt taggaattan taaaaggggg agtatgcctt tccagaatct tccatcttcc 660
tttgganacc tggccttcct cccatttcct ccctttggcc ccaggtanga aggatggagg 720
gaggnttgtt actnttnccc ttctgggggc cctttctggg ggcctaaccc tgncaatttt 780
anttccnccc tcccttacct ngggatgnng ggnccctttn ccgggattta ancctggggg 840
ctgggcccta antttttcc ctttttttc ccnaaaaaaa aaaaaagggg ggggccccc 900
ctgnnnnngn nttttnnaa aatnccccc nngncntnng gnccccnccn nnccccnntt 960
tnnttnancc nccccctgggg ggtcccnttt ngggggnnnt tnnntttnna nccnnnnnnn 1020
ggggnttttt ttttnnnnna aaantttttt ttnnncnnnc nnnnncnnnn nncnntttnn 1080
nnnnggggggg gnggntnnnn nntttnnann ncccccntttt tnngnnnaaa anncccnnnnn 1140
nnnnnggggg gggnnnnnnn nnnnnnnnnn nnnncncccc cnnnnnnnnn nnnnnnnnnn 1200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1320

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1380
nnnncngnnn cnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnc 1493

<210> 19
<211> 1602
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1602)
<223> n = A,T,C or G

<400> 19
ggaaaatcaa gatgtggctg aagatcagag gctcagttag caacctgtgt tgtagcagtg 60
atgtcagtcc attgattgtc tttagagagt taatgttaca aaaaagaatt cttaataatc 120
agacaaacat gatctgctga ggacacatgc gcttttgtag aatttaacat ctggtgtttt 180
tctgaaaaaa tatatataca tatattgctt tatttgaaac aaattaaaat atgctgcatt 240
tgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 300
aaaanaaaaa aaaaaaaaaa angggggggn ccccccccng gnnngnnttt ttgnaaantc 360
ccccccccnn gannntngggn ncccnacnnc ggccccannt ttantttaan cccnnccccc 420
cttggggccc ccctnnnggg ggggntttna tttccaaaan cccccaanng ngggggttnt 480
tnntttcncc aaaaancnnt ttttttnnaa accncccccc ggaacccccn cccccccttt 540
ttcntttaag gggggggggg gnttttnttcc ccccttttтg gaaaancccc ctttttttt 600
tgggggggccc aaaaaaaacc ccccctttng naccnnnnan gggggggggg ggggnaancc 660
tttgggaaaa ccccccccng gggagngaaa ancccctttt ttcccccccc ccctttttgt 720
tttcctnngc cccaaaaacc ccntcccccn ntgggggann tnggcnggng annenannan 780
cccnnaaaan gncccccccc cccnnnnggn gaaaaannec cccnnaangg ggnttntntc 840
ccnggggana aaaancccng gggggggncn ttttcccccg tttngncccc naaanggggg 900
gggcccccct tgggcnnnna aaaaccccct ttnntnccen cccccgnggg ggggnnnttt 960
ccccccnaaa ntcccccccc ctngccccna angggaaaac ccccnnngng gggtcccttn 1020
gggnncccce cnnttttttc cccccccngg gcgggggggng nnggggggga nnccccgnng 1080
gggcctttcc nnnngttttt ccncccennec cctntnnngg gggtgaaann aaccccccen 1140
ngnnttnntn ancccccena nannnngncc ccnnttttтg tnccccccnc cngaanncnn 1200
acccccccce ccnannttтt tttgnnnngg gncncccccn gngnntnntt nnccccccce 1260
ccccccccce ccgggggngn ggnttttttt gnnnnnnnnn nccccengggg ggggngcccc 1320
ncccccncnc ggnntttggg ngnnncccce ctnnttttntt tnnnnccncc ccccccccce 1380
cgccntttnn gnnggnggng nnnnncngcn ccccccctnnn gntcnnttnt cncccencen 1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn cnnnnnnnnn nnnnnntnnc 1500
ncncnngcnn tcnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnngnnngng nnnnnnncnn 1560
nnnncnnncn nnnngcnnnn ngnnnngcnc cgcccncnnn cc 1602

<210> 20
<211> 1633
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1633)
<223> n = A,T,C or G

<400> 20
agcacgccag ccatcagccc ctgaatccac ctcacccact cgccagacct ttttgtcgaa 60
gttcatgtcc ttccttagcc ttccaatgaa gcctctacct gcctgagatg tccaaggtaa 120
tccatcagct gaggctctca gagaatgaaa gtgtggccct gcaggaactc ttggactgga 180
ggagaaagct ctgtgaggaa ggacaagact ggcagcagat cctgcaccac gctgagccca 240

gggtgcctcc cccaccacct tgcaagaagc ccagccttct gaagaagccg gaaggggcct 300
cctgcaacag gctgccgtct gagctctggg acaccaccat ttgatgtggc ctgaactgca 360
gacttacaaa atagaactgc ctactgattc cgggctgcaa caacagaagg ctgccttctg 420
acatgcgctg gggcttctct ccacgcattt agacaaaaaa agcacaggac acagacacta 480
aatatatgag atcccgtgtg tgtgtgtgtg tgtttgtgtg tgtgtgtgtg ggttctttct 540
tatccatctc gnggngatac actctgattt tcaagctcct catttacggg tcttgtgcta 600
cccctaggta ncaagaaaan aggctgggaa aaagtgtggn cgtggncnan agcgananaa 660
gtancggnng gaaaggagcn antccatgca cacttctgta ccngtngttt tttntacngg 720
ntcaaacagg nntgnntnat tggncnttnc caangggggt tttntttant aannaccnng 780
nnntnncngg ggannaanan nannnnnnna nnnnnnnttt nggnnnnccn cccttggggg 840
ggnnnnnantt ggggcncnct cnctcccccc cctcncncccc ccctcccccc tcacnncgnc 900
ncnccntnnn ccncggcgcn nctccncntc nncnnccnnn ntcgncccnn nngnggggggg 960
gcggggnngn ncccccnctct nctccncnnn ccccccccnn cnccnncncn ncnncncccc 1020
cncccncncc nnnncncccc ccncnncccc ncccccccnn nnnngnnnnn nnnnnnnnnn 1080
ncnnncccce ccccccncce ccccccnncn ccnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnggggggcn ngnnnnnnnn nnnnnnnnnn 1200
nnnnnnnncn ncccccccccc cnnnnnnnnn nnnnnnnnnn ncccccccnn nnngnnncnn 1260
nnnnnngnnn ngnggggggn gnnnnnnnnn nnnnnnnnnn nnngnnnnng nnnnnnnnnn 1320
nnnnnnnnng ggnnnnnnnn nnnnnnnnnn nnnnnnnnng nnnnnnngnn nnnnnnnnnn 1380
ngnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngnnnnnnn nnnnnnnnnn nnnnnnngnn 1440
nnnnnnnnnn nccgncccce cgnncnnnnn nnnnngnnnn nnnnnnnnnn nnnnnnnnnn 1500
nnnnnnnnnn nnnnnnnnng gggnnngcgg ngnngngggn nnnnggnnnn nnnnnnnnnc 1560
cnncccccnn nnnnnnnnnn nnnnnnnnnn nnnnngnnnn nngnnnnnng nnnnnnncen 1620
nncccccccg nnn 1633

<210> 21
<211> 1462
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1462)
<223> n = A,T,C or G

<400> 21
gggctcccaa aatggcgaag tgaggctgcg gggactcgct gagcagcgga gggggagcgt 60
gcagagccgc tgcggccctc acagtccgga gcccggccgt gccgtgccgt agggaacatg 120
cacttttcca ttcccgaaac cgagtcccgc agcggggaca gcggcggctc cgcctacgtg 180
gcctataaca ttcacgtgaa tggagtcctg cactgtcggg tgcgctacag ccagctcctg 240
gggctgcacg agcagcttcg gaaggagtat ggggccaatg tgcttcctgc attcccccca 300
aagaagcttt tctctctgac tcctgctgag gtagaacaga ggagagagca gttagagaag 360
tacatgcaag ctgttcggca agacccattg cttgggagca gcgagacttt caacagtttc 420
ctgcgtcggg cacaacagga gacacagcag gtccccacag aggaagtgtc cttggaagtg 480
ctgctcagca acgggcagaa agttctggtc aacgtgctaa cttcagatca gactgaggat 540
gtcctggagg ctgtagctgc aaagctggat cttccagatg acttgattgg atactttagt 600
ctattcttag ttcgagaaaa agaggatgga gccttttctt ttgtacngaa gttgcaanaa 660
tttganctgc cttatgtgtc tgtcaccagc cttcgagtca anantataan atgtgctaag 720
gaaganttat tgggactctc ctatgatnac nattnatgga naacccggtt ggccttnaac 780
cttctttttg ctcanacggt nttaaaatat ttagncgngg ggngggatct ttggtcaccc 840
aaggaaaaan nacccggnaa ntttaaaatt ttttgnnaaa aaaaaaannn ttccnaaaaa 900
gggaatttct ttnaaanttg gccccaaana ccttgnggnn ctttnggnnn ntttgnnctt 960
ttnannnccn nnggggggnng nnttnccnna aaaaaaattt nntttnnngg gnnnnncnnn 1020
nncannnnna annnnnnnnn nnnnncccnc cngngnnnnn nnntnnaaag nnttttnnng 1080
gnncccnnaa aatngggggn ncnntttttt nttttnccnn nnnnnnnnnn nnnnnnnggg 1140
ggggggggnc ccnnnnnttt ttnnnnnann nnnnnnnnnn nnnncnnncc ccnnntnnaa 1200
annnnnnnnn nnnnnnnnnn aannnnnnnn nnnnnnnnnn nngggggggn nnnnnnnnnn 1260
nnnnnnnnnc ccnnnnnnnn ncnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1320

nnnnnnnnnt ntntngnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn gnnnnnnnnn 1380
tnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1440
nnnnnnnnnn nnnnnnaaaa an 1462

<210> 22
<211> 1601
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1601)
<223> n = A,T,C or G

<400> 22
cccgaagcac gacgcagagc ctccggtgtg gctgtctctg atggtgtcat caaggtgttc 60
aacgacatga aggtgcgtaa gtcttcaacg ccagaggagg tgaagaagcg caagaaggcg 120
gtgctcttct gcctgagtga ggacaagaag aacatcatcc tggaggaggg caaggagatc 180
ctggtgggcg atgtgggcca gactgtcgac gacccctacg ccacctttgt caagatgctg 240
ccagataagg actgccgcta tgccctctat gatgcaacct atgagaccaa ggagagcaag 300
aaggaggatc tggtgtttat cttctgggcc cccgagtctg cgccccttaa gagcaaaatg 360
atttatgcca gctccaagga cgccatcaag aagaagctga cagggatcaa gcatgaattg 420
caagcaaact gctacgaaga ggtcaaggac cgctgcaccc tgcagaaaan ctgggggnca 480
gtgcccgtca tctccttgaa ggcaaagcct tttgtgaacc cccttctggc cccctgcctg 540
gaagcatctt ggcaagcccc cccncctgcc ccttgggggg ttgcnaggct tgcccccctt 600
ccttcccana accggaaggg gcttgggggg gatccccccan caggggggga aggggcnant 660
ccctttnccc cccannttgg ccnaaaccng nnccccccnc ncccccttgg nanttttncc 720
nttnttnccc ttcccatncc cntttngcng gggtnnttng gnccctttcc ccnaaannng 780
gggnttttnn gnaancnttt ttnnaaannn nccentnttt gggggnctnn nnaaannccn 840
naancccnna nngtntnncc ccccccccnn ngggncnccc cccccnnnnt nttntnnnng 900
ggggggggnn aaancccccn nnnnnnnnnn nnnnnnnnnn nnaaaaanaa aannantncn 960
ccccccnntt tttcccnccc ncccnncngg gggnnccnnn tccccccccn ttttttcccc 1020
nannnnnnnt gggnnncnna annttttttt tnnancccnn cnnnntnnnn nnnnnnctcn 1080
nngnnnnnnt ttnncnntnt nttnnnnnnn nnnnnnnnnn nnnnnnantn nnnaannnnn 1140
nnnngnnaaa acnatncccc ctcnctttnn ccccnnggnn ncnnnnncct ttnnccccnn 1200
nnnnnnnnnn ttttncengn nnnncnnnaa nggcnccttn nnntnaannn ncccctccc 1260
nngnnnngnn nccccaangg nganaantgg ggnnccccnc ccccnnngcn nnnnaanttt 1320
nnnttnnggg gnnnnnnccc cccgccgcgc ctcccnctcc ccttcgcgcc gccccgcgcc 1380
gccgtccgcc ccgccccccc nctcccnctc cccgccgtcn ctcncttcnc tctccnccgc 1440
gccccgcccg cgcgcccgct cgncgtcncg ncncncnncn ccnnnnnnnn nnncgnnnnn 1500
ananaagnnc nccnaccnat cccccccgcc nnccccccnt nccgnnnnng nnnnnnnnng 1560
nncgcccncc ncccccnncc cccnttcgtn cccccccntt n 1601

<210> 23
<211> 1566
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1566)
<223> n = A,T,C or G

<400> 23
ttttttttt tttttgattt tttttaatgc tgcacaacac aatatttatt tcatttgttt 60
cttttatttc attttatttg tttgctgctg ctgttttatt tatttttact gaaagtgaga 120
gggaactttt gtggcctttt ttccttttc tgtaggccgc cttaagcttt ctaaatttgg 180
aacatctaag caagctgaag ggaaaagggg gtttcgcaaa atcactcggg ggaagggaaa 240

ggttgctttg ttaatcatgc cctatggtgg gtgattaact gcttgtacaa ttaccgtttc 300
acttttaatt aattgtgctt aaggctttaa ttaaatttgg gggttccctt cttagagcag 360
ctcgtactga cgaaggtgca tgcgctgaat gatgtcacgg cagtcgttga acacacggcg 420
gatgttctca gtgtcccagc gcangtgaaa tgagggtagc agtagtgacg cccatctcca 480
ctggcagtgc tgatcctcag aaactcatct cgaatgaagt acttggcccn ggtcacgcgt 540
gggtnctctt cnggctcngg agtancatnc tcangagtag ggtagcgagc aaattctgga 600
aagaagcctc aatcttcnat ttcccnncaa ggactttctc ancganccan atcttgcttg 660
tttganggaa ccaggaatcc cngnnnaatg gngcncaacc ccttcttgtt ggttncccaa 720
aangcccntt gaaaaaaggg ttcaaaaanc cctccctgcc anggccgggg ttngggncct 780
gggnttgncc ccccccccgg naaaaaancn ctnntttnnn naaancttgn nttggnttgg 840
ggnccccccc ccccnaaaaa aaaanaaaag gggnnnnnnn ccncccennt nntttnnaaa 900
aanaccccng gggnannccc ccccttttgg gggggggnnn tnnntttnnn nnncnnnggg 960
ggcccccccc cccnnnnnaa aaanaattnt ggggaaannn nnnannnttt ttnnccccnc 1020
ccnnngnnaa aantnngnnn tnncnnaaaa tnncccnaaa nnnnnngccc ccnnnnnnnn 1080
aaaannnnnn nntnnnnnnn nnnnnaanaa nnnnncccnn tntannncnn nnnnntnncn 1140
naaaanngng gcncnnnann nnnnnnnncn tngnnnnnnn nnnnnnnnnn cnnttttnn 1200
ccnnaanntn nnnnntnnnn nngnggggnn aannngncnn cncccnccna annnncccnc 1260
nnnnggggnn nccccnnngg gcccnnnnnn nnnnccnngn nnnnnnnnnn nnnnnnnnnn 1320
nnnnnnnnnn nncccngnnn nnnnnnnnnn nnnncnnnnn cnnnnnnnnn nnnnnnnnnn 1380
nnnnccnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnngnnncn 1440
nngncnccnc nnnnnnnann nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnncn nnnnnnnncn 1560
nncccc 1566

<210> 24
<211> 651
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(651)
<223> n = A,T,C or G

<400> 24
cgtcggttgg cgactcccgg acgtaggtag tttgttgggc cgggttctga ggccttgctt 60
ctctttactt ttccactcta ggccacgatg ccgcagtacc agacctggga ggagttcagc 120
cgcgctgccg agaagcttta cctcgctgac cctatgaagg cacgtgtggt tctcaaatat 180
aggcattctg atgggaactt gtgtgttaaa gtaacagatg atttagtttg tttggtgtat 240
aaaacagacc aagctcaaga tgtaaagaag attgagaaat tccacagtca actaatgcga 300
cttatggtag ccaaggaagc ccgcaatgtt accatggaaa ctgagtgaat ggtttgaaat 360
gaagactttg tcgtgtactt aggaagtaaa tatcttttat tagagaaagt gttgggacag 420
aaagtacttt atgtaactaa gtgggctgtt cagaacttan aggcattttt tgtaatttct 480
ttttaattac tttananagc tagggatgca aatgttttca gttagaaagc ctttatttac 540
ttttggaaat tgaacaanaa atgctttgtc ttanaactgg agaatatttg atggtaggga 600
aacatgtaat ggttctctgg caaaattgnn tcannatttg aaatgaaann n 651

<210> 25
<211> 676
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(676)
<223> n = A,T,C or G

<400> 25

```
gggggacaga gactcagatg aggacagagt ggtttccaat gtgttcaata gatttaggag      60
cagaaatgca aggggctgca tgacctacca ggacagaact ttccccaatt acagggtgac     120
tcacagccgc attggtgact cacttcaatg tgtcatttcc ggctgctgtg tgtgagcagt     180
tggacacgtg aggggggggt gggtgagaga gacaggcagc ttgnanntnn ttgcttngan     240
ntttcncnta naacccgcna gcgcttnggt agggtnngcn anggatgncn nncnttnttc     300
nnaagncncc ngttcngngt canttgcttg nctcntctaa ctcnnnnnnc ccccnnttnn     360
gtctcctnng ngntcnaccc nntctgnttc ttngntcnng nttgncctcg nnnttnnttc     420
nnngctcngc ncgtntggtg nnntgngnat nannctnanc gngtttntnn attntnnctn     480
ncgtngancn catntgancc ttntnnngnt nttcgnctnn ntcgancgtn ttcngggncn     540
cnccncgnnt ctnnctnncc tcnccctttt ntcntcttgn ttgtggcntn acctnnctcn     600
ttctntgtnt ncnngccttn nngtgnnncn gatagtcnnc cctntttgnn aatatctntn     660
tnntcncccc cctccc                                                     676
```

<210> 26
<211> 657
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(657)
<223> n = A,T,C or G

<400> 26

```
ttttttttttt tttttgctgg gtggtaactc tttatttcat tgtccggaag aaagatggga      60
gtgggaacag ggtggacact gtgcaggctt cagcttccac tccgggcagg attcaggcta     120
tctgggaccg cagggactgc caggtgcaca gccctggctc ccgaggcagg caggcaaggt     180
gacgggactg gaagcccttt tcanagcctt ggaggagctg gtccgtccac aagcaatgag     240
tgccactctg cagtttgcag gggatggata aacagggaaa cactgtgcat tcctcacagc     300
caacagtgta ggtcttggtg aagccccggc gctgagctaa gctcaggctg ttccagggag     360
ccacaaaact gcaggtagtg atgtgcaaga ntccatcctg cagttttcca gcaatganaa     420
actcctcctg cggttgtggg acctggggaa gtatccgcan acctctcctg gcgggggtgt     480
agacnaaccg gatgtcaccg gcatccccta aagnttggaa ccctttatac atcttgggca     540
tcttganctc ataacgctgg tataaggngg ntnggtngac ttttggnngt ccccccaant     600
gcccttgana ccaaggccgn aattncnaaa ggcccctgng gggggggggg acccagn        657
```

<210> 27
<211> 646
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(646)
<223> n = A,T,C or G

<400> 27

```
ggaangctga agaattaaca ntttgactnc taaatgtgat actggntngt anattccctt      60
agagcagaaa ggagaggggc acatattaat ttgtatcgct tttgcttctc tttggtcttt     120
tgtgtcttag aatttggaag tggttcattt ctgttgctgg tatgaggatt tcgaatactt     180
agtaatcgaa aaccatatcc tgtaatttaa taaaaaaaac taaggaagaa aaaaccctcc     240
aattttccca aatgcaatca gtgtaactag gggctgtgtt tctgcattaa aataaatgtt     300
tcangctttg tggtcctgat caaggtcctc attaaaaaat tggagttcac cctagngctt     360
ttcccctctg tgactgggct cntccccccac cnctcttagg tatcgcagtt attatgggnt     420
ncaaatnaag naatangntt nncaaatttn accaaaanaaa gcatttttt cactgcnttn     480
tnattggggg gttggcccaa ccncntcaat ggntcttanc atggntggnt acccgcnacc     540
tttncntnaa cttggngnaa ncnngggcnn tacnnttcct gggggnaaat ngtntccnnc     600
cantccccnc ncntncnanc cgaancnnaa agggnaancn nggggg                    646
```

<210> 28
<211> 407
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(407)
<223> n = A,T,C or G

<400> 28
caagagtctt tgaaataagc ccatttgagc cctggataac aagggataaa gtggagcgga 60
tgcacatcac agacatgaaa ttgcctcacc tgcctggctt agaagacctt ggtattcagg 120
caacaccact ggaactcaag gccattgagg tgctgcggcg tcatcgcact taccgctggc 180
tgtctgctga aattgaggat gtgaagccgg ccaagaccgt caacatttag tgcctcctga 240
gcagctcttg gttttggcgt cttttgggtc ggcccatgtg gtttgagcac ccagccaggc 300
ggtctcttta gaggatcctg tacacagttc cactattaaa acatttcagg ttgaaaaana 360
nnnannnnnn nnnnnnnnnn nanannnnnn nnnnnnnnnn nnnnnng 407

<210> 29
<211> 625
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(625)
<223> n = A,T,C or G

<400> 29
ttttttttnt ttttttttnt ttttttttng gggaccaaat ttctttnttt gaaggaatgg 60
nacaaatcaa acgaacttaa gnggatgttt tggnacaact tattgaaaag gnaaaggaaa 120
ccccaacatg catgcactgn cttggggacc anggaagtca ccccacgggt ntggggaaat 180
tancccnagg nttanctttc attatcactg nntcccangg ngngcttgna aaanaaanat 240
tccncccagc cacattnngg cnctcccatn ttgcncaagt tggncacgtg gncacccaat 300
tctttgaagg ctttcaccng ctnattnaag naangggtct caatgaaanc acaccantgg 360
ggggnatttt tgntnnnngc ccattgggca attcccaana tggctgaatc aaatttttnt 420
nccaaagnca ngcccctcca atggattnaa ancccentnc caatanaaca nnnggntttt 480
ttatcctcca agaaaaattn ggcccntntn gggntggaag gtttnantat tacaagcncc 540
ttcctttaaa tggggaaaaa nttttgnnaa annttaaaac cncntcgcca agntttnaaa 600
agggnaggna ngcngngggt tacnn 625

<210> 30
<211> 643
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(643)
<223> n = A,T,C or G

<400> 30
cttaagaatt ggcccagcct cagatcctgt ctttagcaac cagctaatat ttacccagag 60
gtactgcaat agagtatttc aaaatggaat caggatctgg tgggcctcag aaattgtctc 120
ttttctgagt ttcaatttgg ttctcctgga tgttttgctc tgttttggta cctgtaatat 180
agggaaacac aacttttttt gggaaagccc tttgacccca gcttgctagt tgcataataa 240

taaattttct gttcctaaaa aaaaaaaaaa aaaaaaaaaa aaaanaaaaa aaaaaaaaaa 300
aaaaaaaaaa aaggnngnaa naaaaaaata anangggncc gntaaaacnn gggggggggcc 360
cntcaanttt aaagggccct ttaaancccc tnnnnaancc nccntggncc nttttnnttc 420
ccacctttg gnggnnggnc ccnccccccgg nctttttttg ncctgggggg nccccccccc 480
tggtcnttnc ttanaaaaan nangaanttg cctcccttnt cngaaaangg ntctttttt 540
ttnggggggg gggggggggg ggaannnggg gggggtggg ggaaaaattn nggggnttg 600
ggaaccnggg gcccttgccc ttnngaaaag aaccccntggg ttt 643

<210> 31
<211> 645
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(645)
<223> n = A,T,C or G

<400> 31
gtgaaagctg taaaacacct tttatggaag aaaagaaata aaatgtagtt gtcaagtcta 60
aaaaatagta gcaacgggaa tcataatgaa tacatgcaat gaatttaaaa tgtaaaaatg 120
aatttaaaaa gtaaaaaggg ctctgtggtg taatttttct taactacaag agtctaaata 180
cactgctttt ctttaagagt tcattttaat tagtaacgtc aaacaaaatt attctagata 240
atgagcccta caaattacta ctactagcaa ctgtcatttt ttactcgggc atcctctagg 300
tgtcttacat tctcatttta ttcttacaac gaactcatcc tccagaagga cttcatcctc 360
cagaaggact catcctccag aangactcat cctccaaagg acttctccag aaggggggaaa 420
tggaagaccc gggtaacttg ctcagggctt atcacagaac tatgtttgag cctgacttcg 480
tttgaactct aaagcccaca tgctctttct actgccccat gcttctcaag gnaccagact 540
cttatttnct gcacttttga gaatctnaag atcctgantc attttaaata aatttagttt 600
tttggggagn agccnnaaaa aaaaaaaaag ggcgccctcc ncnnt 645

<210> 32
<211> 668
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(668)
<223> n = A,T,C or G

<400> 32
tcccgttctg ttttaaacag aaaataaaag gagtgtaagc tccttttctc atttcaaagt 60
tgctaccagt gtatgcagta attagaacaa agaanaaaca ttcagtagaa cattttattg 120
cctagttgac aacattgctt gaatgctggt ggttcctatc cctttgacac tacacaattt 180
tctaatatgn gttaatgcta tgtgacaaaa cgccctgatt cctagtgcca aaggttnaac 240
ttaatgtata tacctgaaaa cccatgcatt tgtgctcttt ttttttttta tggngcttga 300
agtaaaacag cccatnctnt gcaagtccat gtatgcngcn cttaagcntt ctatctttgc 360
tcaaatngnt gaangatggg gaccttggct catggcttgc gnatttgatc ntaangnncn 420
tttctancta tgntatgagg cacnngccct attggaggnc gccccnggtt tccggaaaag 480
ngcnntnntg tngngaattg cnnctcggan ttcaanaata tncggcnntt gntttgnang 540
ccnngnnnan caatcaggng ngcccctcna antcatgnaa gccccgnntn aannenctnc 600
nctnttctcg nnntgggnnt tccattgccn gcctcgacgn ggttngcctc tcnccggcnn 660
cncgcncg 668

<210> 33
<211> 682
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(682)
<223> n = A,T,C or G

<400> 33
ggcttgtccg agttgatatg cgtatgcttt gcctaaaaag ccttaggaaa ttagacttga 60
gtcacaacca tataaaaaag cttccagcta caattggaga cctcatacac cttcaagaac 120
ttaacctgaa tgacaatcac ttggagtcat ttagtgtagc cttgtgtcat tctacactcc 180
agaagtcact tcggagtttg gacctcagca agaacaaaat caaggcactc cctgtgcagt 240
tttgccagct ccaggaactt aagaatttaa aacttgacga taatgaattg attcaatttc 300
cttgcaagat aggacaacta ataaaccttc gctttttgtc agcagctcga aataagcttc 360
cattttgcc tagtgaattt agaaatttat cccttgaata cttggatctt tttggaaata 420
cttttgaaca accaaaagtc cttccagtaa taaagctgca agcaccatta actttattgg 480
aatcttctgc acgaaccata ttacataata aggattccat atggctcttc atattcattt 540
ccattccatc tctgcccagn atttggggat acccgcanaa aatttggggt ttgggggggaa 600
aaatntggnc tggaactttt tttanttnaa gggaaataat naggggngga agggggggggt 660
ttntggntgc ccccccccng gn 682

<210> 34
<211> 1549
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1549)
<223> n = A,T,C or G

<400> 34
ttgagagata cctccctcct tctgctcagc tgccttgcag taattaaact ctttctctgc 60
tgcaacaccc ctactgttct ccgtgtattg gcttttctgg gcagcaggaa ggaaaagctg 120
atgcgatgct ctcagtgccg cgtcgccaaa tactgtagtg ctaagtgtca gaaaaaagct 180
tggccagacc acaagcggga atgcaaatgc cttaaaagct gcaaacccag atatcctcca 240
gactccgttc gacttcttgg cagagttgtc ttcaaactta tggatggagc accttcagaa 300
tcagagaagc tttactcatt ttatgatctg gagtcaaata ttaacaaact gactgaagat 360
aagaaagagg gcctcaggca actcgtaatg acatttcaac atttcatgag agaagaaata 420
caggatgcct ctcagctgcc acctgccttt gacctttttg aagcctttgc aaaagtgatc 480
tgcaactctt tcaccatctg taatgcggag atgcaggaag ttggtgttgg cctatatccc 540
agtatctctt tgctcaatca cagctgtgac cccaactgtt cgattgtgtt caatgggccc 600
cacctcttac tgcgagcagt ccgagacatc gaggtgggag aggagctccc atctgctcct 660
ggatatgctg atgaccagtg agggagcgcc cggaagcagc tgagggacca gtactgcttt 720
tgaatgtgac tggtttcccg ttgccaaaac ccaggacaan ggatgctgga tatggcttaa 780
cctgggggga tgaaccaang tttttgggaa ngggaaagnt tnaaanaaaa tcccctggna 840
aaaaaaantt tnnaaanaaa accttggaan ggggccccc ttgggaaaaa nggggggggan 900
nnnggttnt tnggnccnnt ttnnccccn nnnnannnct ttaannnngn nnantttttt 960
nnaangggg nntnnccccn ntttnnaann ntntntcccc nnnnnanggg ggggtnncnc 1020
nnnccccng ggggnncnnn ntnaacnccn nnctntnggn ggaaancntt tttttncttc 1080
nnccnnggnc cccnanannt tttcccagaa nccccccng ggggngnnng gaaangnnnn 1140
nnnccctcnn gggggttncc ccnnnaaaaa aaannnggnt tttttttttna nganccgggg 1200
acnccccnnn naaanntttt tnnaaagcgc cccccnnnnt nnggnnnnnn nggnannnnn 1260
nnnttngnnn nttngcccnc cnttnnnngn nccnctcnnn nnnnnnnnnn nnnnnnnnnn 1320
nnnnnnnnnn nnnnnnnnnn cntntanntn ntgnaaaaaa nggnnnnngn nnnnnnnnnn 1380
nnnnnnnngn cccccnngng nnnnnnnnnn nnnnnnnnnn ggggggnngn ggnnngcnnn 1440
nnnnnnnnnn nnncnnnnnn nnnnnnnnnn nnnnnnnnnn ncgnnnnnnn nnnnnnnnnn 1500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnncngnng nnggnnanc 1549

<210> 35
<211> 1440
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1440)
<223> n = A,T,C or G

<400> 35
ctaatctaag cctcaaactc gttattgggg ctataaagaa aacgtttact tacccagctg 60
aaacaggtta agaatattct taatctcatt atagataatt gcccccatgg gacttgaaat 120
acaacacctt gtgctgaaaa cttcaggttg gcaatatttg aaggtttcgt tgtagaagag 180
tttaacatta actcctattt tgacttacaa atcttgtttc tcatcactaa aatgcttttg 240
aattaataat ccaacccaca tgagctgaga gtttttcttt tgttagaaaa gaaacagaca 300
tctttctgta tgaaagtata aattgtatgg ttttagatac ataagaattg acaaaagcga 360
gcgaaatctt tgtacttctg agttcttgct gtatgtatgt tttgttttaa atctgattag 420
ggacacccag cagctggccg ggattcttgg attgctcctt gggagttaag attgtcaata 480
ctcctgtgaa gcaagggatt tcagccatag aacaaagatt tattgttgcc acctgaaaag 540
tttacaagta tttattgtgt atttgataca ttgcttgaaa aagatgaaat ctgttaaaga 600
ttcttttccg atgtccaggt taagaagaaa cctccttgta ttgagtgaaa ttatatgtta 660
aatggtatta gagaatgtag gtggnataga aattggattt ttcttggngg tngaacaacc 720
tcaagttcgg caaagtttaa aatttggatt aaacaagaaa aannggttca nggttgnaaa 780
angggacttg nttagggang ggacaanggc ctttaaanna ccngcgtccc ttctccnggc 840
nggcnngncg ggcccnnccc caanctnntc cangcnttcg nccncnaccn nccnccttt 900
cctnntnnca cnaanntctt tnnccntttt tacngggggn ggggnnnccn ncnccggcnn 960
cngnntncgc cncccanaaa nnccnncntt ttccnncnnc ccntttncnn nnnctttnnc 1020
cnnnncccc cccgnnnnnn nnnnnnnnnn nnnnnnnnnn nggnnnnnnn cccnnnnnnn 1080
nnnnnnnnnn nnnnnnnnnn nnnnnnggnc nngggnnnnn ttnntnnnnn gggggncnnn 1140
nnnnnnngcg nnnnnnnnnn ngnnnnnnnn nnnnncgnnc nnnnnnnnnn nnnnnnnnnn 1200
nnnnnnnnnn ncccnngnna ncnaannncn nnncnnnnnn nnnnnnnnnn nnnnnnnncn 1260
cnncnnnnnn nnngnnngnn nngnnncnnn nnnnnnnnnn nnnnnnnnnn nnnnnngnn 1320
nnnnnncgnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnng nnnnnnnnnn nnnnnnnnnn 1380
nnnnnnncnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn gnnncgaaga nggccnaccg 1440

<210> 36
<211> 1496
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1496)
<223> n = A,T,C or G

<400> 36
tgcataccgt ggaagggcgc cagggtcttt gtggattgca tgttgacatt gaccgtgaga 60
ttcggcttca aaccaatact gcctttggaa tatgacagaa tcaatagccc agagagctta 120
gtcaaagacg atatcacggt ctaccttaac caaggcactt tcttaagcag aaaatattgt 180
tgaggttacc tttgctgcta aagatccaat cttctaacgc cacaacagca tagcaaatcc 240
taggataatt cacctcctca tttgacaaat cagagctgta attcacttta acaaattacg 300
catttctatc acgttcacta acagcttatg ataagtctgt gtagtcttcc ttttctccag 360
ttctgttacc caatttagat taagtaaagc gtacacaact ggaaagactg ctgtaataac 420
acagccttgt tatttttaag tcctattttg atattaattt ctgattaagt tagtaaataa 480
cacctggatt ctatggagga cctcggtctt catccaagtg gcctgagtat ttcactggca 540
ggttgngaat ttttcttttc ctctttgggg atccaaatga tgatgtgcaa ttcatgttta 600

acttggggaa acttgaaagg ggttcccata tancttcaaa acaaaaacca aatggtgtta 660
tccngacgga tctttttatg ggtnctaact agtactttnc taattgggga aaagnaanng 720
ctttnagttt tgcnnaatta agtttggggg aagggcnata attaaaaatt gagggccccg 780
tnacnaaaac caactggggg ngtntaacga aaaaccctgt tttnaaaagg gggccttttn 840
ccccttnnnn ngnnatntna nttnccccnt ttgccntttc cntttnnnnn naaacttttt 900
nnnttttctc cccnancnnn naaangngna nngggtntcc ccccnangtt nnnnttnttc 960
nnnnnnannna nccccccctt ngnggnnccn nnngggcntt ttctcntngn naanngttnt 1020
nnnannccct tttgncnnnn gggnnttgng nttcggnngn ccnngggggn nnnnccnnnn 1080
gnnngnnnnn gannangann nnnggnggnc gtntnnnngg ccgcgggnnn nngngnnncg 1140
ngnnnnnngn nnnnnngnnn cnnngnnnnn nngnnnnnnn nnnnnangnn nnnnnnnnnn 1200
nnngngnnng ngnnnnnncn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnncn 1260
nncnnntntn aancnnnnnn nnnnnnnnnn nggnnnnnng nnnnnngngn nnnnngnnnn 1320
nnnnnnnnnn nnggnnnnnn nncnnnnnng nngnnngcgg nnnnnngnnn nnngnnnnnn 1380
nnnnnngnng gnnnnnnggn gnnnncnnnn nnnccgcnnn nnngnncncn cnnnnnnnnn 1440
gncnncnnnn cnnnngnnnn nncnnnnnnn nngnnntnng nnnnnccgnn gnnntc 1496

<210> 37
<211> 1604
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(1604)
<223> n = A,T,C or G

<400> 37
atgcagtcct ggatggagcc gactgcatca tgctgtctgg agaaacagcc aaaggggact 60
atcctctgga ggctgtgcgc atgcagcacc tgattgcccg tgaggcagag gctgccatct 120
accacttgca attatttgag gaactccgcc gcctggcgcc cattaccagc gaccccacag 180
aagccaccgc cgtgggtgcc gtggaggcct ccttcaagtg ctgcagtggg gccataatcg 240
tcctcaccaa gtctggcagg tctgctcacc aggtggccag ataccgccca cgtgccccca 300
tcattgctgt gacccggaat ccccagacag ctcgtcaggc ccacctgtac cgtggcatct 360
tccctgtgct gtgcaaggac ccagtccagg aggcctgggc tgaggacgtg gacctccggg 420
tgaactttgc catgaatgtt ggcaaggccc gaggcttctt caagaaggga gatgtggtca 480
ttgtgctgac ccggatggcg ccctgctccg gnttcaccaa caccatgcgt gttgttcctg 540
tgccgngatg gaccccanag cccctccttc agccnctgtg ccacccccct tcccanccaa 600
tccattaagn cannaangct tgtanaactt cactctggnc tgtaaacntg gncacntgtt 660
nggtngggac accttgggaa ggaaaaatca acncctcant tgnaaaattg gggtaangnt 720
tgccantcnt gtttttaaan gggacnagnc gcgaggaagg gctnanttnn ttanantnnn 780
agggggcccc cnnccccnat nnananggggg caaanaacgg nanggnaaat ngnttnnnnc 840
cttngnnngc nccccccnnng gannnccccn nncgnggnnn nnnnagnggg gntcancnnc 900
ntnccccttnt nctnnntgng gtnnnccnnn nnnccnnnnn cacgttnaaa annnaaatnn 960
ngncccnnnn gnnngcctca cncnnttngn ggnnngaccn anccaccnng cnnatnggng 1020
ntggnagggn ctctncnnca aancantnng gncttcgtna ngngtgnnnn nnnnnnnnna 1080
ncnngntnnn nncncnnngc nannnngtnn cnngnntccn cccacttgtn tnncnannng 1140
ngtnnnngnn tngannntcn nngnttgnat cccggnaana cnannnncgg ncncnggcnn 1200
nccnncnncn gnncnntccc nnnccccnatn nnggnggnnn nctgcnanct nnnnngancn 1260
cnnnnnnnnn gncncanncg antngngnng nnnntnncnn nnnnnnnnnn nnnnnnnnnn 1320
nntnnnnnnn nccgnttntg ctngcagtac tntcgngnnt ntcnnnnnnn ngnnnnnnnn 1380
ncnnnnnnnn nctngnacnt tngnacgcnn nagtcgacnt nctnggacnt nntnnncant 1440
cnngccnngt nnngnntngn ngcnnacnnn nnnacnnngg cgnnnnnnnc ncatnncnnc 1500
nctnanannn ggtnngngng nnncccttccn nnnnagnnnn natanngncn nnanncnccn 1560
nnnnnnnnnc ngnnnnncnn nntcnncgaa nanntgncac nacg 1604

<210> 38
<211> 280
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(280)
<223> n = A,T,C or G

<400> 38
ttttttttt tttttaattt atcagngctt aaaaatcttc aaaatagctt agtgaggctc 60
atgacagtgc tggccccatg gaaatgtagc cttttgttgc gtttaaacac tgtcacacca 120
tctatgactg tcccattggt ctgaagtgta gtggcaaact aagcatccta taagacaagc 180
taaagcttgc tttttgccag tcagttgaaa gtcttgcatc tcttcactga tgcactttct 240
ttaggtattg atagtcagaa gcacaaagca tttattatgc 280

<210> 39
<211> 378
<212> DNA
<213> Homo sapien

<400> 39
cgagtttata atcctataat gaagaatact ggcacaggca atgctcactc gaaaacttca 60
agtaatttct agttggtttt ggaatgcttg ataaagttcc tttacagctt tattttcctg 120
atttgttttg gtttagatca aagttcaaat taattttaac ttagctaatg aactcatcac 180
caggacagtt ggaggggtga ggccgaggtt aaatggtcca cgtttcaaaa atgttaatgg 240
ctaatccata attaaagaag gtttaactgt tactgaagtt tacaagtttt attgtcatga 300
acatgaaata caaacacgat ggcttcgaaa tgtctttcaa taaatgtttc tgcatttata 360
tggaaaaaaa aaaaaaaa 378

<210> 40
<211> 2039
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(2039)
<223> n = A,T,C or G

<400> 40
caacttttgt agaagtattt ttttctctgt aatattttta ttggctcata aagatgtttt 60
catatctgaa ctcctaaata agtgaaatta cagtagatta tattaacaaa atactttta 120
ggtagccatg cttgagactt tttaaaaata taacttttc cttaaagttt tcagctatag 180
caaaaggtag ttatgtatgc cagacctaat atgagctgcc accaacaccc ctagaacttt 240
cagccatggt gtcttcagaa ttgtagcgca tttctgaatc tagcaaatcc tccttttacc 300
cgttgaatgt tttgaatgcc ctgactctac cagcgcccat aaatgatctc tagaaggact 360
gttagtacca acctgttttt caactttgaa gctaaaaacc ctgatatggt aatattatgg 420
tgcatagcag aggtctcgga aaaaaaatat ttctgttcac tttactttca ggttaaaaat 480
gtttctaaca cgcttgcaac ttccccttatg gcattaatct tgttgaggga gagagacaga 540
atcctggact ctccaaagta tttaactgaa agtagggcct gctctgacag ggcccatgtc 600
ccacaaggct ggcttnggcc tcaggggggg gctttggctg gtgcttggga tgaaaattgn 660
tggannncngg tntttgggga taaanggacc aaanggacca gccaaaagcn aaaaaatngg 720
gnttttaaa ngccttgggg ggnttacctt tttcntttaa angnnggttt naaagnatta 780
gggctaaang ccantttnca aaaaaangct cccnananaa aatggtggaa aagggnccct 840
tttggncgac aggncctttg nggaaaattg cccccancng ggccctttt tgnccccccc 900
nncccaaaaa aaagntgggn ngaagnnttn ttaaaaccct nnnggngccc ntttttttng 960
nnaaanccnc cnccnnggggg gncgccccnc tttnttttntt ntnttcccng ggngnccnnt 1020
ttttttncgg cngacccnnc ggggntcaan nnctgnanaa gnngntatct ggcngggnnn 1080
gcgcnngaaa gnnnnnggnn ncngnggggg nnnncgcncg nnannnttnt gnggggnaaa 1140

aaaaaaganc cctctnttnc tctcttntnt naanntnnnn ngnnnnnnan ncnngcnnnn 1200
gnngngnngn nnnnnnngnc nnncnnannn ggggggnggg cncncncnnc nnnnantnng 1260
gggcgnctcn tnnnnnnccc cnctncgggn nccnnnncnn ggngngngcn nntntngnng 1320
tccngnntgt gtntgnnnng ncnnncncnc cncgnnnnnc tnnnnntntg nntnngnnng 1380
ggggngnncn nncccncncg tgnnnnntnt nnnnnnnnnn nnganggnna nnncnnncnn 1440
nnnnnnnnnn ggggngcnnn nncnnncnnn tnnnnnnngg gnggnggggn gnnnnnnnnn 1500
nnnggnnnng nnnnnnnnnn nnncncncnn nnnnnntgng cgnnnnnncn nncnnngnnn 1560
nnnnntnnnn nnnnnnnnnn nnnncnnnnn nnnnnnnnnn nnnnnnnnng nnnnnnnnnn 1620
nnnncnnnnn nnnnnnngng gnnnnancgn tgngcngnng tnnnnnnnnn nnnnnnnnnn 1680
nnnnnnnnnn nnnnnnngnn nnnnnnnnnn nnangnnnnn nnnnngnnnn nncnnnnnnn 1740
gnnnnnnnnn cnntgcgagc nnnngncnnn nncnnntgnn nnnnnnngnn tcgcncnnnn 1800
nnnnncgngg ggcgntnnnn nccncccgcn gntgncnnnn nngncnnnnn ncnnnnnnnn 1860
ngnnntnncn cnnnnnnncg nnnnnnnnnc nnnagngnnn ngngnncnnc nncnnnatnn 1920
gannnnnnnn nccnncnnnn nnnnncgnnn nngcnnngnn ngnnnnnnnn nnnnntcncn 1980
ncncnnngnn nnngnnnnnn nnncncncgn gngnnnngnn cccgtccgcg cgngcgcgg 2039

<210> 41
<211> 319
<212> DNA
<213> Homo sapien

<400> 41
ttttttttt aaaaaaaaag agtttattta gaaagtatca tagtgtaaac aaacaaattg 60
taccactttg attttcttgg aatacaagac tcgtgatgca aagctgaagt tgtgtgtaca 120
agactcttga cagttgtgct tctctaggag gttgggtttt tttaaaaaaa gaattatctg 180
tgaaccatac gtgattaata aagatttcct ttaaggcaga ggctggtcga gatgctgctg 240
ttatcttctg cctcagacag acagtataag tggtcttgtt tctaagattc ctaccaccag 300
ttactttggg ccaagtatc 319

<210> 42
<211> 524
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(524)
<223> n = A,T,C or G

<400> 42
ccttttttnt ttttttnttt ttttctgatt tcaagtcaag atttattgct ttacaaacaa 60
acattatact tggtcttaat agaaaaatga caccagatac atccaaaata catttcacat 120
tgggatagct gccagttcag cacaaaacat acattactag gagcagggag gcatgaaaat 180
aaactatatc ttacttttttg gtacgtcagg aacacttttg cctgaagtaa gcccttttagt 240
actatttttt attttattta ttttttttaat ccacccatct gcacactggn cctttagtac 300
tctttaagta taaaacttta cttgtcctgg gctttgaccc ttgtgtttga tctaaatgac 360
atttcaaaca taaatgtctt ttgactagtg cgcttactgn tatgtacana atttaaaatg 420
tgatcgttng aatntaaaat ctggtttgat acatgatata aaagttgtat atttaaaatn 480
caagaaatgt ttttggggaa tatttctact aaagaatttt aaat 524

<210> 43
<211> 103
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(103)

<223> n = A,T,C or G

<400> 43
ccttttttt ttttttttgc nngaaataag gaatctataa atctgaaata aagaaatccc 60
attttaaatt aaattgttaa agagacacat aagaaaaaac act 103

<210> 44
<211> 425
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(425)
<223> n = A,T,C or G

<400> 44
gtcgacaaga taatgtactg acatctctag caatcttttt tgccagtggc tttaaattgc 60
caataagtta aagaatattg ttcctatggg ttaaattttt attcttattt tcacatttaa 120
atttatttt cttaattttt gtggatacat aatatgtgta tatatgtatg ccatatatgg 180
tatattttga tgcaggcata ctctatataa taatcacatt agaggaaatg agatatccat 240
tacctctagc atttattctt tttattacaa gncaattcaa ttgtacactt tttagttatt 300
tttaaattta caatgttatt gattacaggg tcatttttat ggtcataata aaaaatttta 360
tacaaaacgt gtaaaatcta tacatttctg agttctgaat aaatattttt taaaaatttt 420
aaaaa 425

<210> 45
<211> 492
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(492)
<223> n = A,T,C or G

<400> 45
gtcgactgcc cccaccgctg ggcggcgctg cggggcaccc aggctctgca gtcagcgccg 60
cgccgggaat cctgtacccg ggcgggaata agtaccagac cattgacaac taccagccgt 120
acccgtgcgc agaggacgag gagtgcggca ctgatgagta ctgcgctagt cccacccgcg 180
gaggggacgc aggcgtgcaa atctgtctcg cctgcaggaa gcgccgaaaa cgctgcatgc 240
gtcacgctat gtgctgcccc gggaattact gcaaaaatgg aatatgtgtg tcttctgatc 300
aaaatcattt ccgaggagaa attgaggaaa ccatcactga aagctttggt aatgatcata 360
gcaccttgga tgggtattcc agaagaacca ccttgtcttc aaaaatgtat cacaccaaag 420
gacaagaagg ttctgtttgt ctccggtcat cagactgtgc ctcangattg tgttgtgcta 480
gacacttctg gt 492

<210> 46
<211> 499
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(499)
<223> n = A,T,C or G

<400> 46

ccttttttttt ttttttttat aacatttata taatgtgcta acaatgaatc catccatgat 60
ttattgtttg taatgaactt aaaataaccc tttacaaatt aaaatcattt tttcaaacat 120
gacttcatat tgaaatggtt ctgttaaaaa agtaaaagtt gaattttcca gccaatttag 180
catctaggac ctgaatcttg ccaatatcct acccactatc ttcattccta cctcctaccc 240
cttcaaatca gctcctccag actttcctat ttctgtcacc ccagttcaaa atggttttca 300
ccatgcattt gatgtaaaat gtgcaagtgc gatatgactt cacaaagtat caattgtgtg 360
gacaatgata actactgtga cactgctagc acccctggct aaaagtaaga agcaacaaaa 420
ttacacaggg ttccttttctg atgaatgcag nanggattca agaaatccca ganctggaaa 480
aagattttca atagatctg 499

<210> 47
<211> 537
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(537)
<223> n = A,T,C or G

<400> 47
gtcgacattt ttctgaggaa tagtttgtga ttccaatgca ggtgtcttca ttaccattac 60
ctctacactg cagaagaagc aaaactcctt tattagaatt actgcacatg tgtatgggga 120
aaatagttct gaaaggctag aatgatacaa gtgagcaaaa gttggtcagc ttggctatgg 180
agtggtggca ataatctcta aacattccaa aagaccatga gctgaaccta aactcccttg 240
gaatctgaac aaaggaatat aaaattgcca tttgaaaact gaccagctaa tctggacctc 300
agagatagat cagccagtgg cccaaagcca tttcaagtac agaaattata gagactacag 360
ctaaataaat ttgaacatta aatataattt taccactttt tgtctttata agcatatttg 420
taaactcaga actgagcaga agtgacttta ctttctcaag tttgatactg agttgactgn 480
ttcccttatc cctcaccctt tccccttccc tttcctaagg caatagtgca caactta 537

<210> 48
<211> 556
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(556)
<223> n = A,T,C or G

<400> 48
gtcgactttt ttttttttt ttagnnntat aaaatatttt atttacagta gagctttaca 60
aaaatagtct taaattaata caaatccctt ttgcaatata acttatatga ctatcttctc 120
aaaaacgtga cattcgatta taacacataa actacattta tagttgttaa gtcaccttgt 180
agtataaata tgttttcatc tttttttttgt aataaggtac ataccaataa caatgaacaa 240
tggacaacaa atcttatttt gttattcttc caatgtaaaa ttcatctctg gccaaaacaa 300
aattaaccaa agaaaagtaa aacaattgtc cctctgttca acaatacagt ccttttaat 360
tatttgagag tttatctgac agagacacag cattaaactg aaagcaccat ggcataaagt 420
ctagtaacat tatcctcaaa agctttttcc aatgnctttc ctncaactgn ttattcagta 480
tttggccagt acaaaataaa gattgggtct caactctctc tttcattagt ctcaagngtt 540
cctattatgc actgag 556

<210> 49
<211> 355
<212> DNA
<213> Homo sapien

<400> 49
gtcgaccgag cctctcccac cctcagtcgc atagacttat gtgttttgct aaaattcagg 60
tattactgaa ttagcgttta atccacttcc tttcttcttc ttctaaaata ttgggcactc 120
ggttatcttt taaaattcac acagaaaaat tccgtttggt agactccttc caatgaaatc 180
tcaggaataa ttaaactcta gggggacttt cttaaaaata actagaggga cctattttcc 240
tctttttat gttttagact gtagattatt tattaaaatt ctttaataat aggaaaaggg 300
gaaagtattt attgtacatt attttcatag attaaataaa tgtctttata atacc 355

<210> 50
<211> 507
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(507)
<223> n = A,T,C or G

<400> 50
ccttttttt ttttttttaa aaaaaaaaaa ttctgtttat tgtaataatt aaataagagt 60
aaacatttta aaacatataa aaataacttt aaaatatagt aacactttac aaaatatgta 120
tctaattaaa aatacattaa catagcatcc ctcaaactat acaaatatag aatatatatt 180
catgaaattc tttanaaata taacatctat tctttgaata aagcttaaaa tttgtttata 240
attttcaaac taanaaaaga agtagngaat aatagctcca tccaatttat aattgtctta 300
aagagaatga ttatgtatca tttcttgctt gtcttttcta atacccagtc aatcacctgt 360
acagcattgt tgtttgctgt tttcttcatt tcttcaaata gacccccttga aagtttttaa 420
gatcctttag atagaactta gagatttcaa agagacgctg gctgcatgca gtgaaacatt 480
catgagtctc ggtaatactg ngtttct 507

<210> 51
<211> 538
<212> DNA
<213> Homo sapien

<400> 51
gtcgacgcaa aagtttgact aaactttacc tttttatagt ttcacttttt aagttatatt 60
tagaatatat tgatagatta taaattgatt gtgaaacttt tttctgaatt ttttcaacat 120
gttttactca gttacatgag ttaaaggata ttttcagtcc tgttatcttc aattgcagtc 180
tttaaaaaaa cccaccctat tgttctactt gttatatgtc tattcataca gtaaattcat 240
ttcaaggttt atgccagtgg gtattattgg tgctttttga agttgaggtg aaccatccag 300
gaaggtcttg ttaatgttat gttcatctat aatggcatag gggaaatata tatattttta 360
atattgtaaa catttgtact gaataacctt tttttccccc cctccgcaag caaaactggt 420
tgaacagcgg atgaagatat ggaattcaaa gctctaatgg acctttttga agagaagttg 480
tggcttatgt ggagtttaca tgggcctctg atggaagaaa gctaatctgt ttagtatt 538

<210> 52
<211> 504
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(504)
<223> n = A,T,C or G

<400> 52
ccttttttt ttttttttta aagtacaaat tcagtttatt catctgttta tgacacagta 60
cacaggaggc aaagtgtttc acatcataga cttcacttcc aactccttgg aatgttcatt 120

tctttggctt acaggagaga ctagacagga aggccaggca atgcttaggc aactaaaatg 180
aggttggggg taatgctaac gtcaccctca cagggatggc cacggggact gttattcgca 240
agctggtttt ctagacctgt tagctggaag catggtgagc accatttctg gacgctcagg 300
ccgtntcggg cttcagtcat ntccaccaca caggtacagc agcgctttct ggtagtcgcc 360
cttagtgtct tgctggatat aatagtacag ggacttgccg tactttctct tgaattcaga 420
cctaattttc aacatgtcca cttcactgng ggagaccatg attctgatca ggacccttat 480
ctcgcgtccc cttgcccttc atgg 504

<210> 53
<211> 489
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(489)
<223> n = A,T,C or G

<400> 53
gtcgacttta gatgtacagg ctgacanana agattcccga gagtaaatca tctttccaat 60
ccagaggaac aagcatgtct ctctgccaag atccatctaa actggagtga tgttagcaga 120
cccagcttag agttcttctt tctttcttaa gccctttgct ctggaggaag ttctccagct 180
tcagctcaac tcacagcttc tccaagcatc accctgggag tttcctgagg gttttctcat 240
aaatgagggc tgcacattgc ctgttctgct tcgaagtatt caataccgct cagtatttta 300
aatgaagtga ttctannatt tggtttggga tcaatnggaa agcatatgca gccaaccaag 360
atgcaaatgt tttgaaatga tatgaccaaa attttaagta ggaaagtcac ccaaacactt 420
ctgctttcac ttaagtgtct ggcccgnaat actgtaggaa caagcatgat cttgntactg 480
tgatatttt 489

<210> 54
<211> 577
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(577)
<223> n = A,T,C or G

<400> 54
ccttttttttt ttttttttttt aagaactcaa tacatggctt ttaattattg tctataattt 60
aaggaaataa tcacctacaa ataggatgtt tctcaagttg gcttacaaat ttgttacttg 120
gcagactgaa aacatttccc acagaacaaa tattatacac aatggttggg ttcctttggt 180
taatgcataa tgtttactcc ataatttatt tacccacaaa catgaattga acatttcttt 240
gtgccanaaa ctattctaac actagaaata caatagtaat gaacaaatag aaaaaaatcc 300
tattgtcatt ggtattacat ccatagtttt ttctccaaga gaataaaagt aagtaaaata 360
tatagaatta tagataatga tatatgctat ggtgaaaaac aaagctgggt aaagggatag 420
agaatggggg aaggataatt ttaactgatt attagtagaa tgtactagta tctctgttct 480
aaaaggattt aagataggta ttacttaccg aacctaagta ttacaaataa aatagcaatg 540
cttacactag gaaagacttt caactgagaa gcattat 577

<210> 55
<211> 483
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

<222> (1)...(483)
<223> n = A,T,C or G

<400> 55
cctttttttt ttttttcac caataattat tttattcagg gagtaaatgt tattaattgc 60
caaaatacga attttaaatt tgagaagtac agatttgtaa gtatatattt gtttgaatag 120
tatcanattg gccttttatt ggcttattgg tatttagngc cagcacttac aatgtgaact 180
cagcaacaga agataattct tatgaaatca acattcaact tacatgaaat aacttaaaaa 240
cttaccaaca atagtctaat gattatatac ctttaccaaa caatgtctaa tgaaagtcca 300
aatgtaaaaa tttaaaaatt aaaattatag aatataattt ttacacatca attgttttgt 360
agcaccatct cgcaaagnaa atatcatgtt tattctgtag ctaaaatttc tccccacaag 420
cagaaattgt ttggaatata caaaaagaca acccattaac aagtaacttt aagtaatgta 480
gtt 483

<210> 56
<211> 521
<212> DNA
<213> Homo sapien

<400> 56
gtcgaccaga cttaagcatc gagtttttac catcttccac tttaagctaa gttatgatac 60
ctattccatt cacaattggt gttcttttta aggtttgcaa atttcagcca attttgtagc 120
taagattgtt ctgatcagct caaaaagatt tggcttagtg ttttcattgc aaattataat 180
tgctgtagag ccacacacaa cttttgaact tttaattata agtgttatgg ctaaagttat 240
ttactgaaaa tttcagtaaa atgtgtgaat gtttctttat gtattaacct catagcagta 300
aatgacttgc tgttgtttaa tttttctaag gcatcttaat agacttctgt tgaaaacttc 360
agtgttaaca tttttatagt ttgtactaaa tttaaccgtg atataaaaat gaattttatg 420
catagatcag gaattttaaa ttaaaggttt tttctttaaa aaaaaaaaaa aaaaagggcg 480
gccgctcgag tctagagggc ccgtttaaac ccgctgatca g 521

<210> 57
<211> 542
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(542)
<223> n = A,T,C or G

<400> 57
cctttttttt ttttttaca acttcacatt ctttaatgtt cattcagaat attaaatgcc 60
attaattgac catcattatt ataaaattta ctatttagat aagtgagttt tagtacagtg 120
ctatttaaag tatggaactg ttactggtgt gtgatcagta cagaaattga gactaagcat 180
ttagaaacct agagcaattt gacgtagcaa tcttctgtct gttgaatcta ataacaaaaa 240
aaatttttc aattttgcat atctttttaa aatttaattt gtcaaggaat tcatttttag 300
catattttac aaaaacatca ttctcctatg gagactattt ggaaatacaa ataagaaaac 360
tggttcttac cacagatagt ttttagaaac ctgttttagn gtaaagccat catttagtat 420
aaagncatct attattactg ttactctgaa gtggttactg agcattacaa cagtnggtng 480
gattataagt ttgtttacta aanatgctag gatttattaa ctcatgtata tatttattga 540
ga 542

<210> 58
<211> 261
<212> DNA
<213> Homo sapien

<400> 58

gtcgacagag aaggtctatg tcaacagagt tgttatctca tagagccagt tttcaaagct 60
ccttctgcat tgtcactcac tgatcaggtg atgaattctt cctagatagt cgcccactcc 120
acctcctact taacctgaga ctcattattt agctatttct gcttttgtaa aaataattca 180
gatattaaac tccaatttta atctatcatc caagggtaga tgtagttgct tagtagcatt 240
ttggaaaaaa aaaaaaaaaa g 261

<210> 59
<211> 480
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(480)
<223> n = A,T,C or G

<400> 59
cctttttttt ttttttaaa atatagaagt tctgagttag acctgtttag ctcanaatag 60
tgggctaaac taccataaaa ttctctgtat atcttaaatg gtaatgggtc aaaaactcca 120
gaaaatcatc agttgataac acacctacag ataagtgcat gggtaggagg ggatagccaa 180
gtgcccatga taatttgacc tcagtaaatt aaactgggca atacacatat ttgctattct 240
gatactgcat tagacttata aaattccatc taataagcat tcataaaact ggacctctct 300
gtatatatct agcttagaca gggataggga aaagaataac tgaagaaact agcttacaat 360
agctaggttt cgtcaggctt attctatcca gccagaaacc accaccagag agaagctgag 420
ccattcagct gnctgtctcc tctccctctg tttgaatagt catgcctagg ccttgctgca 480

<210> 60
<211> 493
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(493)
<223> n = A,T,C or G

<400> 60
cctttttttt ttttttggt ccttctgttt atttcatttt ggatactcag tgaatgttaa 60
ttaaccagga aacttaaaag ttatttcaat tatgaacctc ttcaatcctt catcaattat 120
tttgagtatt ctggtcttaa aaacatctct ttcttctaca aacttctgaa agagatgaac 180
acctccacct acaccaaaat aatgtgcttt gctggccaaa agtacacgtc cattttact 240
taacagtcta aggaaagtct ggtgcaaatt actataataa tctgggttgt aaatggtttc 300
tgaggtgaga atgagatcat attttacaaa aagtttttca ctacttagta caagcttaca 360
aaactcagac cactcaccag aaaaaaatcg gcatttatat agttgngtta cttttggttt 420
cctgcatctt ttcacatctg gctcatttac atcattttct tcatcttcca aagtggagtt 480
agctactaca tta 493

<210> 61
<211> 532
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(532)
<223> n = A,T,C or G

<400> 61

ttttttttttt ttttttttgaa aaatataaaa ttttaataaa ggctacatct cttaattaca 60
ataattattg taccaagtaa ttttccttaa atgaactctt tataatgcat aatttacagt 120
ataagtagaa caaaatgtca tgacaaaagt cattgagtac aagacttgta ataaaaaggc 180
ataaaatata tttatacata aacccctttc aaaaaacaag ggaaagcttg agccctcaat 240
atagggcgac acacggagcg ggtgaccgtg caggtacagg tactgtactg atttaaagtc 300
aagcactaga gatagnggat taatactctt ttgccgtaca ctatatacag atgtatagta 360
caagtaacaa tggcaaacag aatgtacaga ttaacttaac acaaaaaccc gaacatcaaa 420
atgaaggtgt gtggaggaaa ggtgctgctg ggtctcccta caactgttca tttctttgng 480
gggcaggggg tagttcctga atggctgngg tccaatgact aatgtaaaac aa 532

<210> 62
<211> 567
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(567)
<223> n = A,T,C or G

<400> 62
gtcgactttt ttttttttt taagtatttt aggcatattt aataaataac ttcagtaaat 60
agcactgtaa aaagtgaact gttaaaacta aaggcactta aaacaagaat gtgactagtg 120
tgaaacaaga tgggcaactc aaatggtgag aagtaaacat acagtggtct gttatggcac 180
taactcaaag taagactcgc gtaggtgaga gctgttgcat agccacagta taacttcaca 240
tgttcattaa aaaggcaaat tgaccgctaa aacttcaaag aaaaagtact cataaaaaaa 300
gtcttacccc aaaattgcaa acaaatacat taaaagatta gaagaggtga tagaaagcac 360
cagacattaa acaaaataaa aataataaaa taaattcaac tcaaaaggtc cccattcagc 420
aaatactttg taaaagtatg gcctgtatgt aaatagttgc taaatcaagg acttttagc 480
agaaaattgc tcggttcttt tatctaaggc ttgaatttgt aaagngaagg cataaaagtt 540
nccaaacatt aagtaactct taaaatg 567

<210> 63
<211> 247
<212> DNA
<213> Homo sapien

<400> 63
gtcgacaaac aaacttggct tgataatcat ttgggcagct tgggtaagta cgcaacttac 60
ttttccacca aagaactgtc agcagctgcc tgcttttctg tgatgtatgt atcctgttga 120
cttttccaga aatttttttaa gagtttgagt tactattgaa tttaatcaga ctttctgatt 180
aaagggtttt ctttcttttt taataaaaca catctgtctg gtgtggtatg aaaaaaaaaa 240
aaaaaag 247

<210> 64
<211> 330
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(330)
<223> n = A,T,C or G

<400> 64
ccttttttttt ttttttttttt tttttgacat ggagtcttac tctgtcaccc aggctggagt 60
gcagtagtgc aagctcggct cactgcaacc tcaggcagga ctatttttaa ttatttttaa 120
tacctgcaaa agggaatctg cacatgcaca tccgtgtttc tacanaaatc tgcgatcgat 180

ggcagatctg tttgcctttg ngtgtccaca tgaaccattt ggcaaaggca tccaatgcta 240
acggggccca ccaactacaa cggaggcaac aactctgngg attttnnttc acagaaagag 300
taaaatttca ttcaaccgtt ccatgtcgac 330

<210> 65
<211> 486
<212> DNA
<213> Homo sapien

<400> 65
ccttttttnt ttttttnact aggcaaagaa ctttattaat ctttgtttca aacttgattc 60
ccaggcttct tcggcttaat tagctgcaaa gaatgaattg tgtataagca aaaactgaaa 120
agagctgcag tgtccaaggg gcttgggctt aaaaatatta gagatctaga ttttatcaga 180
tccataaaca aaaatttctt aaaaagcagt cataatataa aatagcagct cccagtaact 240
tcttcaggtt ttatcttcag aagttgactc aattcagttt gcctcattct tggaagcctc 300
atcaaaattc tccacaagat ctggaacttc atcatcatca tcctctccag tagcaagtgg 360
tgcttttcca tccacagatt gtttgggcag agcttcggcc agtctcctta aactagtcag 420
actatccgca ccaagctggt ttaagatgct gggtagcatt tctgtcagct gctttgtctc 480
agcatg 486

<210> 66
<211> 503
<212> DNA
<213> Homo sapien

<400> 66
gtcgaccgtc agacagcaac tcagagaata accagagaac aaccagattg aaacaatgga 60
ggatctttgt gtggcaaaca cactctttgc cctcaattta ttcaagcatc tggcaaaagc 120
aagccccacc cagaacctct tcctctcccc atggagcatc tcgtccacca tggccatggt 180
ctacatgggc tccaggggca gcaccgaaga ccagatggcc aaggtgcttc agtttaatga 240
agtgggagcc aatgcagtta cccccatgac tccagagaac tttaccagct gtgggttcat 300
gcagcagatc cagaagggta gttatcctga tgcgattttg caggcacaag ctgcagataa 360
aatccattca tccttccgct ctctcagctc tgcaatcaat gcatccacag ggaattattt 420
actggaaagt gtcaataagc tgtttggtga gaagtctgcg agcttccggg aagaatatat 480
tcgactctgt cagaaatatt act 503

<210> 67
<211> 519
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(519)
<223> n = A,T,C or G

<400> 67
ccttttttnt ttttttgaa taaatttttt ttttatttt acaccataat ccaattctag 60
ttatcttaat tgaatttgaa aacttttca attgcattaa atttacaaaa aagttctccc 120
acattacact aaagcattcc tcatgtttca cttccagtac tcagatactg aatgagtaaa 180
atcattttat tggctctctt ttaattaact ccttcaaatg cacattgttt aaaaactgac 240
taggtcaaaa atagttacnc ctgcaggttg acctattcag actttgccaa actcctccaa 300
gttcaatata aattgacgtt ttcagagtac aaagtcaatt ttacggaaac gctgttcctc 360
cttttccatg gagccaatct gggtaatttt ttcattaaaa ttcttcttct gcctgtttgc 420
tgcggaactc tttgagctgc tgtagccgct cgatagtttc anaaatggtg cgttccccgt 480
ggaccttatt gtcctcttgt gcggatatna acagtgcca 519

<210> 68

<211> 495
<212> DNA
<213> Homo sapien

<400> 68
gtcgactaaa gctgaagaga taaaagaggt tgtggggcta tgtcttaaga caaaagaaca 60
tttagaaaac ctcaggaaat gatcagagtg ggatagatgt tactagaaga aacaaagaaa 120
ttgaattcaa ttaggagtta gaatcattta caaagcaatg gggaaagtaa gcccctaaaa 180
actattgtag catatagtaa ccagagccaa actctcataa tatattcccc aaggcaaaag 240
aaaaatattt acaagattgg cgttgtttta tatgtttgca aacttattta ataagtctgg 300
ctttgtagat ttcatatctg agtctgcatt caatcaaaat gtcttggcta aacttcatga 360
aaaaacccca gcctcataaa ttagtagttg gaaaaaggag gcatatttag agctttttca 420
gataattgta tttctttgat acattagact ggacacacag tagtttgttt aaggttaatt 480
gcaatattgc aatga 495

<210> 69
<211> 525
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(525)
<223> n = A,T,C or G

<400> 69
gtcgacgcca ccatgttcga ggcgcgcctg gtccagggct ccatcctcaa gaaggtgttg 60
gaggcactca aggacctcat caacgaggcc tgctgggata ttagctccag cggtgtaaac 120
ctgcagagca tggactcgtc ccacgtctct ttggtgcagc tcaccctgcg gtctgagggc 180
ttcgacacct accgctgcga ccgcaacctg gccatgggcg tgaacctcac cagtatgtcc 240
aaaatactaa aatgcgccgg caatgaagat atcattacac taagggccga agataacgcg 300
gataccttgg cgctagtatt tgaagcacca aaccaggaga aagtttcaga ctatgaaatg 360
aagttgatgg atttagatgt tgaacaactt ggaattccag aacaggagta cagctgtgta 420
gtaaagatgc cttctggtga atttgcacgt atatgccgag atctcagcca tattggagat 480
gctgntgtaa tttcctgtgc aaaagacgga gtgaaatttt ctgca 525

<210> 70
<211> 511
<212> DNA
<213> Homo sapien

<400> 70
gtcgacattt tatatataat actactaatg gcatagatta acaaaatatt ttacatgtag 60
gaaaggacat aagattactt ttaaagaata gtatgaaata cacaatattc aaatgtgttt 120
gcaatgccta ccaaatttca aatgtgcctg gatcatgtat aaattaagga aagaaaaaag 180
gatcatgtat aaattaagga aagaaaaaat gtaagtatac aacctacacg gtaaaaacaa 240
aaaccaaaca cctggttaaa aatatctatt taagctcgag tgtataacct taaacaattt 300
gtgtatcact agaaaaatgg atttattagt aaaatttagg gcagagattt tattttggac 360
accactgcct ttgtagaaaa atccaaagtg gcataaaaag aaaaataaaa tattaaaaga 420
aaaaatatat attatcattc ccatgttccc atcctgttac tagcattgct gttctggtgc 480
atcaatcctg agtactctaa cttttgattt a 511

<210> 71
<211> 464
<212> DNA
<213> Homo sapien

<400> 71

ccttttttttt ttttttttgga agagcttctt gcactgttat aagaaagaac atgtgggaga 60
ttgcaaacaa agcaacataa agagtataca gcctgtagga gtctgactaa agtaaaaaaa 120
actcatgtct ttgtttagtg agtatctgta tactaagtta atgcaatgcc aattagattc 180
aaattaaatc aagtacaagc aaatgtactg aaagtattag gaatgcatca tctactttgc 240
taaataattt gcactccgca ttctgcaatt acatgagcat gccattggta taatattggt 300
tatataacat ttaacatgtt agtttttaaa agaatgtaga tacattcata gagatcagta 360
tttttacaga tgtttttact ataaaaggaa ccatgtataa cattgatttt taccttcagt 420
tttgataata ggctgaagac tgccttcaat cactttaatt tttg 464

<210> 72
<211> 234
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(234)
<223> n = A,T,C or G

<400> 72
aataaaannt gaacaaaagg aaaaggtgga tataaagtgg aacctgtggg aaagaggcaa 60
gggctgcagg acagaagaga ctgggaactg caggggccct gggactcagg aggagatgct 120
gattcagctc ataggtgacc cagtcctggc cccggctgtt cccaagagaa ggctgtaagt 180
acccagggag gtggtaagca ggatggagga aaaatcagag gactgggggt cgac 234

<210> 73
<211> 143
<212> DNA
<213> Homo sapien

<400> 73
gtcgactaaa taagtcaatt cctggaattt gaaagagcaa ataaagacct gagaaccttc 60
cagaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 120
aaaaaaaaaa aaaaaaaaag ggg 143

<210> 74
<211> 533
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(533)
<223> n = A,T,C or G

<400> 74
gtcgacataa tctaggcatg aagagcaaaa atatcccttc cggagtcttt gaagctgaaa 60
atataaaaca aataaaaaat aaaaaaataa aaacccacaa aaatgttgaa ccaaacctcc 120
ctgctaatct ccatgcccac gttctttccc accctgttcc cagtcttctg acaaactgtg 180
tacatagcgg actcctcctt tctcctccga ggtggtttta aaggcttttt ggtgtataga 240
agtttgtcca tttgtaaaac tccggattgc gttcctcccc gccttccgcc ccttcccttc 300
cctaaagtga tgggctttct cttttctctt tttagtttac ccggtttctt tttaagtaat 360
gtggaagaaa atggtttatt ttgtattgng gtattgaata ttgngttcct ttttatgagg 420
caaacctgat tgtaaacttc atgtaactat agactggaaa aaaatgagcc gngccaaaag 480
tctnccccttc tgtttcttca gcacattgac ccatnncaca cacatacaca cca 533

<210> 75
<211> 485

<212> DNA
<213> Homo sapien

<400> 75
gtcgaccttc cctaggctgt ttctgctggg cgctccgcga agatgcagct caagccgatg 60
gagatcaacc ccgagatgct gaacaaagtg ctgtcccggc tgggggtcgc cggccagtgg 120
cgcttcgtgg acgtgctggg gctggaagag gagtctctgg gctcggtgcc agcgcctgcc 180
tgcgcgctgc tgctgctgtt tcccctcacg gcccagcatg agaacttcag gaaaaagcag 240
attgaagagc tgaagggaca agaagttagt cctaaagtgt acttcatgaa gcagaccatt 300
gggaattcct gtggcacaat cggacttatt cacgcagtgg ccaataatca agacaaactg 360
ggatttgagg atggatcagt tctgaaacag tttctttctg aaacagagaa aatgtcccct 420
gaagacagag caaaatgctt tgaaaagaat gaggccatac aggcagccca tgatgccgtg 480
gcaca 485

<210> 76
<211> 417
<212> DNA
<213> Homo sapien

<400> 76
cacgctggtt ttgcatcttc aggagacgct cgtagccctc gcgcttctcc tcggccaatt 60
cgcggaagaa gtggctcacg ccttccagag ccacatcatc gcggtcgaaa tagaagccca 120
gagagaggta ggtgtaggag gcctgcaggt acaaattgac caggctgttg acggctgcct 180
ccacgtcggt ggaataattc tgacgaatct gggagctcat ggttggttgg caagaaggag 240
ctaaccacaa aaacggtgct ggcaggtccc agaagcagga gatggccgag aagatggtcc 300
cggaggttgc aagcggagag gaaatcggag ggcggtcgga ggctggaaga gagtccccgg 360
atctgttccg tccaaacact gttgaagcaa gagacagacc cgcgggaccg cgtcgac 417

<210> 77
<211> 547
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(547)
<223> n = A,T,C or G

<400> 77
gtcgaccttt tattaagaat atattttatc aggcattttg ataacaaact gttactctaa 60
gtataggtga tttacccagt gtattttaaa aagtaaatga atcccactgt agtttttctt 120
gaaggaaaaa tcatttctcc agttgctgag gggtactaaa agcttcatac acattagcag 180
caaagtcttt cacttgctcc attgtcaaca gatcctgaac aaaatgacta ggtgtttcac 240
tgcaaactga atggatctgt ccgtttacta ttggaattat cttagctaaa ggcaggctga 300
cactggaaag actattcata gagttaccat gttgcaggtc ctgttcagta ggtcgaaaga 360
actcagccat attgtctaga agtctactaa aacctcggtt taaacaggta ttcaaaactg 420
tactaaaatc tgggctttcc aacatgtctc tagtttcatt gagaagttta atagtggtaa 480
tgtctcgagg agaangtcca caggcctgca ctgctaatgg agtttcttca tctggcatca 540
tataatg 547

<210> 78
<211> 499
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(499)

<223> n = A,T,C or G

<400> 78
ccttttttttt ttttttttt tttnnaaaaa aaatcttttt ttatttcaaa gattgcttct 60
tatattgaag ctcatattaa agcaacagta caatgttcat aaaatataag tgtgatgccg 120
taacattttc ttacatgtca gaatactgat atttatatgt atactaaaat aagaacttta 180
aaattgtaca aatagataca ttaaaaatga catagaaata gggcgtctnt cactgaaaca 240
agacagttat atctggcacg tattagttta agatgaaagt agaagcaaaa agatttacaa 300
gaatcagcag taacaagatt gatgctcaag agacataatt gtacattgna ttgtacatac 360
attgtatggg tttaagctgg ctgaatntta tatatttcaa gtttaaaaat gcactacata 420
tagagtgtcc agagtttaag gcgaaattac agctcanaac tgntgncctt tctaattttg 480
gggaagcttn tttgacaac 499

<210> 79
<211> 370
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(370)
<223> n = A,T,C or G

<400> 79
ccttttttttt ttttttttt ttttaaggag caatgacatt tcctagaagt tactttaaga 60
atttccctag agggtcgggt atcatctcan ccagatcttt ctcatccttc aaggccctgt 120
ttggtacagc ttgctaggaa gctgttccag actgcagcag ccctctctgg ggtctctcta 180
ccacttccca ggcactcana acttgtgcct cannanactg ttttgtggca ctgncccatt 240
ctctgattct ccatgtgagc tggttttatc ccatccagca tggctgtgaa atcctaaagg 300
ttcaaacccc agccactctt cacctatatt tcccccaaat ggctagcacg ggaaagggcc 360
caaaggtagg 370

<210> 80
<211> 428
<212> DNA
<213> Homo sapien

<400> 80
gtcgacaaaa agggaaggaa ggagagacag ataactctca gtcatttaaa aaactacaat 60
aaaatattat gaattatcaa ttagatcaaa gttcctcaca gctatattta tataggtaaa 120
aaaaaattaa ataggctaaa tgcccaaaaa tttaagactg gcaaaatata cttggctaaa 180
tactgtgcgt ctctattaaa taccatgttt cagaagaatt attaatgaca tgagaatatg 240
ctcaaaatac atattgatat gtgcaaatac atattgcaaa gtaagattat agaatgatcc 300
tagttcaaaa atgtcacata tatatgtatt taaaaaaaaa ggcagttaag atttacaaca 360
aaatgttagt ggtgggacct tctggtagga atacagattt ttttttattc agaagttttt 420
tgatgtcg 428

<210> 81
<211> 533
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(533)
<223> n = A,T,C or G

<400> 81

ccttttttttt ttttttttatt tttaaaattt ttttattttg aaataattat aaattatcag 60
aaagttgcaa acaaagccca gtcaggtccc atgtaccagt ttcactgcca ccatctttaa 120
aggaggatta gacgaatctg actgctaaaa gtggcccagg gattctggag aaaatccaac 180
aggtttgcta tcaggaaagc aatttcactt acaattcagg tttgactgca agtgaaagtg 240
gttgaaacaa gtgagaagnt gattgcttcc tcatataata gtctaaatgt aggtgtccaa 300
gcctggaata gaggtcctgg tcctctaagt tctcaggaac acaggcttct tttagccact 360
ccacatctct agggtgttgt cctcatggtc caaaatggng actggaattc cagccatcac 420
atntgctttc caggcagcaa aatggaagaa ggggcacana agaacagaga tgacaatagg 480
tataaacaag ctctcttttt aaaggagatt cccaggagct gctacatgac act 533

<210> 82
<211> 493
<212> DNA
<213> Homo sapien

<400> 82
gtcgacccgc gaagatgcag ctcaagccga tggagatcaa ccccgagatg ctgaacaaag 60
tgctgtcccg gctgggggtc gccggccagt ggcgcttcgt ggacgtgctg gggctggaag 120
aggagtctct gggctcggtg ccagcgcctg cctgcgcgct gctgctgctg tttcccctca 180
cggcccagca tgagaacttc aggaaaaagc agattgaaga gctgaaggga caagaagtta 240
gtcctaaagt gtacttcatg aagcagacca ttgggaattc ctgtggcaca atcggactta 300
ttcacgcagt ggccaataat caagacaaac tgggatttga ggatggatca gttctgaaac 360
agtttctttc tgaaacagag aaaatgtccc ctgaagacag agcaaaatgc tttgaaaaga 420
atgaggccat acaggcagcc catgatgccg tggcacagga aggccaatgt cggggtagat 480
gacaaggtga att 493

<210> 83
<211> 501
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(501)
<223> n = A,T,C or G

<400> 83
ccttttttttt ttttttgta ataaagacac tgcttttatt tagtttgata tgtttcttta 60
cagaatgcag aaaacacatc ttaaaatcat atagaaggaa ataaaaacac atcagtggtt 120
ggtgaacact tgaatgtgag attggctctc catctcacag agtccaacgg ccatcaccag 180
cccagcgctc aggggagcag gctgcctgca aaggcattgt tgctgttgtt attctgttca 240
ctgccccatc gcctccagtt gctatggcaa caggccattc tgggccagcc acgtctctgc 300
atggcagtgc ccaatggtgg agttgctagg ggcgacggag ctgtttggaa ggcctttcaa 360
agccctcacc tggaacattg ggaattgttt atttttgat gaggncatca gaaataatct 420
tcaccaggtc agatcccact tgtgctcctg tctctggggc accaggggaa actctgactt 480
ggaggcatga gcccagtcac c 501

<210> 84
<211> 454
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(454)
<223> n = A,T,C or G

<400> 84

ccttttttttt ttttttttt ttttatgcta ataaaacatc ataatttaag gactacactg 60
catttttaa ttccataaat tataatcctt taacatatat gaaagtttca tattcttaaa 120
gngctttaaa atatatttaa ttttttaac aagtggaaaa gaatgtttct taaaagacat 180
ttaatttttt agtggaaatt aatattacca aaaacattct gtgcataaca atttgaataa 240
caatttttt atcttcaaga aatgggattt ttatataaaa tacacatgta gcactgaatg 300
ccaaagtgat gggtatccat ggtcanaatt caaaattaga ttcgctatta aacctgtctg 360
gtttgtgtcc tgagtgaana atgatctcga gctggggagg gaggtgcatt gggtaatcag 420
tgcttttgaa ggtgaatttc cttgctgnga aata 454

<210> 85
<211> 509
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(509)
<223> n = A,T,C or G

<400> 85
gtcgaccgct ctcagctctc ggcgcacggc ccagcttcct tcaaaatgtc tactgttcac 60
gaaatcctgt gcaagctcag cttggagggt gatcactcta caccccaag tgcatatggg 120
tctgtcaaag cctatactaa ctttgatgct gagcgggatg ctttgaacat tgaaacagcc 180
atcaagacca aaggtgtgga tgaggtcacc attgtcaaca ttttgaccaa ccgcagcaat 240
gcacagagac aggatattgc cttcgcctac cagagaagga ccaaaaagga acttgcatca 300
gcactgaagt cagccttatc tggccacctg gagacggtga ttttgggcct attgaagaca 360
cctgctcagt atgacgcttc tgagctaaaa gcttccatga aggggctggg aaccgacgag 420
gactctctca ttgagatcat ctgctccaga accaaccagg agctgcagga aattaacaga 480
gtctacaang aaatgtacaa gactgatct 509

<210> 86
<211> 520
<212> DNA
<213> Homo sapien

<400> 86
gtcgacgggc gccagggtct ttgtggattg catgttgaca ttgaccgtga gattcggctt 60
caaaccaata ctgcctttgg aatatgacag aatcaatagc ccagagagct tagtcaaaga 120
cgatatcacg gtctacctta accaaggcac tttcttaagc agaaaatatt gttgaggtta 180
cctttgctgc taaagatcca atcttctaac gccacaacag catagcaaat cctaggataa 240
ttcacctcct catttgacaa atcagagctg taattcactt taacaaatta cgcatttcta 300
tcacgttcac taacagctta tgataagtct gtgtagtctt ccttttctcc agttctgtta 360
cccaatttag attagtaaag cgtacacaac tggaaagact gctgtaataa cacagccttg 420
ttatttttaa gtcctatttt gatattaatt tctgattagt tagtaaataa cacctggatt 480
ctatggagga cctcggtctt catccaagtg gcctgagtat 520

<210> 87
<211> 171
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(171)
<223> n = A,T,C or G

<400> 87
gtcgacgagt acagtatcag ctgagctgac cttactctga ggactaactc ttttgctgga 60

agcggtttct gatttacagc tcttggtttc tcccagacat gttggtggga gagattttgg 120
tttttaaggg gttgttagat ggagtaaann ttctttaagn nttaattttt t 171

<210> 88
<211> 508
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(508)
<223> n = A,T,C or G

<400> 88
ccttttttttt ttttttttt tttttgnagt aaaaaatctt tatttccaaa atgatttgtt 60
agccaaaaga actataaacc acctaacaag actttggtaa gaaagagact tgatgcttct 120
tataaattcc ccattgcaaa caaaaaataa caatccaaca agagtcatgt tacccattct 180
tagccattaa cctggtttta agtctccaaa atcaggattt taaaatgtac ccaactggga 240
ccaaatacaa acatgagaca ctagggnggc ttgtccttga ttaggaatca ccagcttaag 300
gaactttatc atgggctgag agttagatag atagcttana acaacattgc aaaagngggt 360
gcttctacat gaggactttt ttccccccaa gtagaaaaat aattaaatct tgngtttctt 420
tatattgngc ttttttggg agaaagcaat tcatttaagg atttaaaaca tgttggatac 480
aaaggtagtt canagatgta ataatggt 508

<210> 89
<211> 508
<212> DNA
<213> Homo sapien

<400> 89
gtcgacggga taaatagaaa gcagaatgaa ttaatggaaa agaactcggc tgttaggcca 60
ttctctaaat tctagtttag ccaaaagttt atgtgtggtt tggggcttca tttatttatc 120
tcatgagtaa aatggaataa tacctaacag gcaggctctg gaagttggaa atcacataca 180
cacacacaca cacacagaca cacacacaca cgatcaatca tgtagctcat attagatgtt 240
caataaataa cagctactac agatgcctat cagttgagta agtagttcat taaattgagc 300
tcccaaaggt ctcttctctt cacatccata tccgtttctg cagcaatcaa atagatacat 360
gattgttttt ctgtaagaaa ttactgcaaa gagaatcttt ttctcctact aactgttcct 420
tctacctggt ataggagata aatgtacgtt tcttaattag ctgacttttt agtatgtcat 480
ttctgaagga aaaataaatt aaccttaa 508

<210> 90
<211> 531
<212> DNA
<213> Homo sapien

<400> 90
gtcgacacga gtcccgcgtt ctctccttga atccactcgc cagcccgccg ccctctgccg 60
ccgcaccctg cacacccgcc cctctcctgt gccaggaact tgctactacc agcaccatgc 120
cctaccaata tccagcactg accccggagc agaagaagga gctgtctgac atcgctcacc 180
gcatcgtggc acctggcaag ggcatcctgg ctgcagatga gtccactggg agcattgcca 240
agcggctgca gtccattggc accgagaaca ccgaggagaa ccggcgcttc taccgccagc 300
tgctgctgac agctgacgac cgcgtgaacc cctgcattgg gggtgtcatc ctcttccatg 360
agacactcta ccagaaggcg gatgatgggc gtcccttccc ccaagttatc aaatccaagg 420
gcggtgttgt gggcatcaag gtagacaagg gcgtggtccc cctggcaggg acaaatggcg 480
agactaccac ccaagggttg gatgggctgt ctgagcgctg tgcccagtac a 531

<210> 91
<211> 426

<212> DNA
<213> Homo sapien

<400> 91
gtcgacaatt gaggcctaca agagagggga gcctaggagc ttggattgac cttctagtca 60
accacctgac ttcagcacac cattacaatc gggagactaa accaacaacc agaggatcta 120
aaatgtcaca ttcagatttt caggaagaaa atcttcatta cagtggagca caaatgttcc 180
atacaagaca tcattgagga gccatgctgt ccccttctaa cctgaaacac attctttccc 240
atcctggttg ggcttctgta cctccttatt aatttatgaa cctgaagttg cttgaagtgt 300
tttgggctta ataaatgggg tgaaagtata ggtagcagta acacctacat gaaacaatac 360
accttggatc ttttaatcta aattactttt ctttttaag tctactttta aaataaatac 420
ttctgt 426

<210> 92
<211> 223
<212> DNA
<213> Homo sapien

<400> 92
gtcgactttt aaagcaattg actaggagaa actatttgta gcttatataa caaggactat 60
atataaataa aaaactattt ctatgaaaat cttaaaatta cacacagtcc gatgaaaata 120
atcatatatt aaaaaggcaa accagaaaaa taaatacaga tgaccaaaat ccatgtgaca 180
tatttggcct aattagtaat tagaaaaaaa aaaaaaaaaa aaa 223

<210> 93
<211> 486
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(486)
<223> n = A,T,C or G

<400> 93
ccttttttt ttttttttt ttttttttt tctcaaatat ccaattttat tttatcattc 60
tcgcattggg ggatgcgatc tgcagctagg atcggaattc ccaggcctat anatttttaa 120
accacaccac aggggtaaac cttaaaagaa gngaaaccta acactatata tatttccatt 180
tctaaataca gtatattaca naagtttaaa tatnccacct ntgngtactt acaactntaa 240
aaagatncaa tanctctacc aattataaat aatgtancat ttcatattaa agacattatc 300
gtncaatgga anaataggaa ccctntaacg tatcactatc aaggttagng tctatatcta 360
cttganataa aatactgaaa attcagngta tgaagccaaa tcctgattta acaagttatt 420
ggtagtataa gtgataagtg ttanctgatg aagggaaggc aaatgtggta atttatatct 480
ctgaca 486

<210> 94
<211> 214
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(214)
<223> n = A,T,C or G

<400> 94
ccttttttt ttttttttt ttttttttt tttttngcaa cacaagtcaa tctttattga 60
aaactgcagt attaatacat aacaattctt gttacaataa acgtgctttt ganattttta 120

aatctgagct catctcatca gattgcataa aaaattaaaa tagtntcaat tgacacctaa 180
ctgaactggc tcaggatgga aattccattc cttg 214

<210> 95
<211> 463
<212> DNA
<213> Homo sapien

<400> 95
gtcgaccaga attcagagcg aatggtcaca gttggtcgct gggcaaaggg aatgagtgca 60
gactatgaag aaattttgga tgtacctaaa ccgcaaaaac ccaaaacaaa aatacctaaa 120
gttgttaatt tttgataaca gctagcacta tcatgagtta ctacctcatt gttactttct 180
aaaccaggcc cgcttcacga gttagagttg agctcccctg tagccaggac tatgctgtag 240
atatcagtat gatctgggtg tggccaaaaa caattttctt tattctgtct atcaaatagt 300
acttctacca ctgtttggag aaaattgaag aaaagaataa gatgattaaa tgaattctct 360
aaaagaacat attttaagag acagaactta gacataacca agtagttgta tacctgattg 420
taacaatcat cttttataaa agcaaaatta tgcataaatg taa 463

<210> 96
<211> 606
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(606)
<223> n = A,T,C or G

<400> 96
gtcgacttta aaagtgcctc ggcatcctgt attacatgtc atagaattgt aaagtcaaca 60
tcaattacta gtaatcattc tgcactcact gggtgcatag catggttaga ggggctagag 120
atggacagtc atcaactggc ggatatagcg gtacatatga tccttagcca ccagggcaca 180
agcttaccag tagacaatac agacagagct tttgttgagc tgtaactgag ctatggaata 240
gcttctttga tgtacctctt tgccttaaat tgctttttag ttctaagatt gtagaatgat 300
cctttcaaat tgtaatcttt tctaacagag atattttaat atacttgctt tcttaaaaaa 360
caaaaaaact actgtcagta ttaatactga gccagactgg catctacaga tttcagatct 420
atcattttat tgattcttaa gcttgtatta aaaactaggc aatatcatca tggatacata 480
ggagaagaca catttacaat cattcattgg gccttttatc tgtctatcca tccatcatca 540
tttgaggcct aatatatgcc aagtactcac atggtatgca ttgngacata aaaaagactg 600
tctata 606

<210> 97
<211> 530
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(530)
<223> n = A,T,C or G

<400> 97
ccttttttnt ttttttgta gatttttgc tatgttactc aggctggtct tggactcctg 60
ggctcaagcg atcctcccac cttggcttcc caaagtgcca ggattatagg catgagccac 120
catgctcggc ctgctccttt tcttgaaaca cctcctctgt ggtttagatt ccaggagact 180
ggaatggtct gccctggtgg gctgctgagt cagggacctg aggtgtttgt tcactgggga 240
ggcgggttca gatcaggaat gtaaggatga tggaaagaag ggagtcactc tggtttggtg 300
ggactgggga gcaatcttga tcacggccac ttacagcttc tgccattgtc cttcaccact 360

atctcagcat ctcggtccct cacgatgtcc ctccagtcaa ttgtgtccat gtgacaaagc 420
ttatcgttct tctcaatata aacacccccct gacagaatct cggtgagctg agtcaagcgg 480
agctggcgca nagcgtggct ggagttggtg ttatagttca acatgacgaa 530

<210> 98
<211> 479
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(479)
<223> n = A,T,C or G

<400> 98
gtcgacggtt agtttctgcg acttgtgttg ggactgctga taggaagatg tcttcaggaa 60
atgctaaaat tgggcaccct gcccccaact tcaaagccac agctgttatg ccagatggtc 120
agtttaaaga tatcagcctg tctgactaca aaggaaaata tgttgtgttc ttcttttacc 180
ctcttgactt caccttgtg tgccccacgg agatcattgc tttcagtgat agggcagaag 240
aatttaagaa actcaactgc caagtgattg gtgcttctgt ggattctcac ttctgtcatc 300
tagcatgggt caatacacct aagaaacaag gaggactggg acccatgaac attcctttgg 360
tatcagaccc gaagcgcacc attgctcang attatggggt cttaaaggct gatgaaggca 420
tctcgttcag ggggcctttt tatcattgat gataagggta ttcttcggca gatcactgt 479

<210> 99
<211> 502
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(502)
<223> n = A,T,C or G

<400> 99
ccttttttt ttttttgta agtttaaatt tatttttaa aaatgcttgt cttcctcact 60
agacaatcaa ctctatgagg gcagagacta tgtcaccact gtcccaccag cccctggcac 120
acagtaggta ctcaataaat atatgttgga aggatggatg gaggtaatgg atggaaagat 180
ggatggaagg atgaatggag ggatggatgt gacccagctg aagtgtgagt aggaacattc 240
tcttattatg ggtggaggaa agagagagga gattgagaaa ataagataaa atacattgat 300
gagcatcatt tttggtgttc gaaaagtagg attgaattag gactaataaa tctagagaat 360
tttacctctt tcaatgccca agccacactt ttctatcact ttgaaaccga aaaagtaaat 420
actttcccaa catttgcttt gctggtagga aatgctttaa taaaaatgca atctctangt 480
tgccatggca tcattaaaag aa 502

<210> 100
<211> 537
<212> DNA
<213> Homo sapien

<400> 100
gtcgaccctt tccataaatc cttgatgatt gacaacaccc atttttcctt ttgccgaccc 60
caagagtttt gggagttgta gttaatcatc aagagaattt ggggcttcca agttgttcgg 120
gccaaggacc tgagacctga agggttgact ttacccattt gggtgggagt gttgagcatc 180
tgtccccctt tagatctctg aagccacaaa taggatgctt gggaagactc ctagctgtcc 240
tttttcctct ccacacagtg ctcaaggcca gcttatagtc atatatatca cccagacata 300
aaggaaaaga cacatttttt aggaaatgtt tttaataaaa gaaaattaca aaaaaaaatt 360
ttaaagaccc ctaacccttt gtgtgctctc cattctgctc cttccccatc gttgccccca 420

tttctgaggt gcactgggag gctccccttc tatttggggc ttgatgactt ttctttttgt 480
agctggggct ttgatgttcc tttccagtgt catttctcat ccacataccc tgacctg 537

<210> 101
<211> 611
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(611)
<223> n = A,T,C or G

<400> 101
gtcgacctaa aatgaagtgt ttgaaatcag aaatctattt ctaatgtctc atagctttaa 60
aactattttt gtccttatac tcatacttgt tattttattt tattcatcct atatagccat 120
ttgactgaaa tgtagaaaat aatttattaa attgagaaaa tatgcaggca ttgaacaatc 180
tttcaagtat tttgaataaa aattcaaatt attatagatt gcctggaatt gttaagactg 240
tcagaaggtc agctcattga tagctaagta gtatacactc tgaaaaacag aatgtagaaa 300
tgggttttat aaaagctgac ctctagagta aaggaggacc cagcatgtgt aattcttcct 360
cttaatactt taagaccact aatttgagga cttatggttt ctcaccactg cactcttgca 420
gctttcaaga aagtacttaa gttttaaatg cccaggtgat ttctaagact cttgaataga 480
attggttggg ttcttctgat attgcatttt catgagaaaa aatttcagtg gtacattaat 540
ttttattttt ccttttgctt atagacttcg catatcattt aaagtgatgg ttcgagcttn 600
ctctggatac t 611

<210> 102
<211> 498
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(498)
<223> n = A,T,C or G

<400> 102
ccttttttttt ttttttttta acgcatattt gtttttattt ataggtaact accacatgaa 60
ttataaagac aacaaaggat gtcagaatga acatggatag gtgtatgcat actacggcta 120
aggagaaaca atgttcctac atattatggg tagtgagaac attatctgta taacagggaa 180
ctgtgattat ttaaaaatat gcagaactta tttcatctgt gctttanaaa taactgtata 240
cagtgttata agttgaaaag aactcaaaat aactaatacc aaatatacac ctatgtatta 300
naattcaaaa aagctgcttt ctgtgaagtc aatcagctat attaaaaaat gacacaaatc 360
caaaacaaga tgcatgttat atataaaggg acattgtaag tttccttgct gcattaaacc 420
catggtttaa tccatgaaat ttccttttaa ttatcattta gacagaagca tgcaaatagt 480
ctcaggatct acttaaga 498

<210> 103
<211> 446
<212> DNA
<213> Homo sapien

<400> 103
gtcgactctt ggtgtttttg tatttccacc tcacccccag cacatagccc agtctcttgc 60
acaaattaag tacttaatgt gtgttgagct aaattgaata aaggattatt agcattagca 120
tattttgtgc cttggttgta taagctggtt gtttgttttg ttacctttgc aaatatttat 180
gattatcacc cccccacata ctaaattgtt tttaaaagtt ttgcctttcc ttcagatact 240
accccaggca atttgctgta gataatgtga ttgcttccaa tgacataatt atcccaaact 300

ctctgccccg gatatacttt gccaaacgaa atttgaattc tctgaataaa ttggtcatgt 360
cctaaaaaaa aaaaaaaaaa aaaaaaaggg gcggccgctc gagtctagag ggccccgttt 420
taaaccccgc tgatcagcct cgactg 446

<210> 104
<211> 286
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(286)
<223> n = A,T,C or G

<400> 104
gtcgaccttc gttatccgcg atgcgtntcc tggcagctac attcctgctc ctggcgctca 60
gcaccgctgc ccaggccgaa ccggtgcagt tcaaggactg cggttctgtg gatggagtta 120
taaaggaagt gaatgtgagc ccatgcccca cccaaccctg ccagctgagc aaaggacagt 180
cttacagcgt caatgtcacc ttcaccagca atattcagtc taaaagcagc aaggccgtgg 240
tgcatggcat cctgatgggc gtcccanttc cctttcccat tcctga 286

<210> 105
<211> 406
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(406)
<223> n = A,T,C or G

<400> 105
gtcgacgcgt agcagagtgg tcgttgtctt tctaggtctc agccggtcgt cgcgacgttc 60
gcccgctcgc tctgaggctc ctgaagccga aaccagctag actttcctcc ttcccgcctg 120
cctgtagcgg cgttgttgcc actccgccac catgttcgag gcgcgcctgg tccagggctc 180
catcctcaag aaggtgttgg aggcactcaa ggacctcatc aacgaggcct gctgggatat 240
tagctccagc ggtgtaaacc tgcagagcat ggactcgtcc cacgtctctt tggtgcagct 300
caccctgcgg tctgagggct tcgacaccta ccgctgcgac cgcaacctgg ccatgggcgt 360
gaacctcacc agtatgtnca aaatactaaa atgcgccggc aatgaa 406

<210> 106
<211> 258
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(258)
<223> n = A,T,C or G

<400> 106
gtcgacgatt ttttttgtac attttggctg cagtattggt ggtagaatat actataatat 60
ggatcatctc tacttctgta tttatttatt tattactaga cctcaaccac agtcttcttt 120
ttccccttcc acctctcttt gcctgtagga tgtactgtat gtagtcatgc actttgtatt 180
aatatattan aaatctacag atctgttttg nacttttttat actgttggat acttataatc 240
aaaactttta ctagggta 258

<210> 107

<211> 369
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(369)
<223> n = A,T,C or G

<400> 107
gtcgacgtaa aatagaaaca gaaggggact ttatcaacct gattaacttt ctcaacatgt 60
taaccctaca gttaacatta taatcaatgg tgaatcattg agtactttcc ttctaagatc 120
agaaacagtt caaagtccac tctcaccatt tctattcaac attgtactgg aatcccagcc 180
agtgcagtaa taccaataat aaaaaattaa agtcataaag attgaaaagg atgaagtaaa 240
gctatttcaa ttntatttag aagtatttag aaaccccaaa gaatctacaa aaaactaata 300
gaaataagtg aatatatgaa ggtcttacta tacaagatca acatatcaaa agcagtggta 360
tttaagaaa 369

<210> 108
<211> 289
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(289)
<223> n = A,T,C or G

<400> 108
gtcgacattg catccttgaa atcctgggct caggtgatcc tcccgcctga gcctcctgag 60
tatctgggac tacagatgcg tgccaccaag cctggctaat tttgtctcat gtcttctaaa 120
aattattttg tgaagcccct tcacaaaaaa ccttaaggga aatctgatgg tgctcaggaa 180
tctaactctc cctaaaccat cctctttaac tgcttctaaa atatctctgt tggcctttct 240
tancctttt ctgtttccat tcagtgctcc aagcgctttt tgtttctaa 289

<210> 109
<211> 444
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(444)
<223> n = A,T,C or G

<400> 109
gtcgacctgg cgttggcacc gctgaggaat gggcctgggc ggggagggac atctctacac 60
cgttcccatc cgggaacagg gcaacatcta caagcccaac aacaaggcca tggcagacga 120
gctgagcgag aagcaagtgt acgacgcgca caccaaggag atcgacctgg tcaaccgcga 180
ccctaaacac ctcaacgatg acgtggtcaa gattgacttt gaagatgtga ttgcagaacc 240
agaagggaca cacagttttg acggcatttg gaaggccagc ttcaccacct tcactgtgac 300
naaatactgg ttttaccgct tgctgtctgc cctctttggc atcccgatgg cactcatctg 360
gggcatttaa cttcgccatt ctctctttcc tgcacatntg ggcagttgta accatgcatt 420
aagagcttcc tgattgagat tcag 444

<210> 110
<211> 196
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(196)
<223> n = A,T,C or G

<400> 110
cctttttttt ttttttcatt aaataancca tcatcacatt agtacaatac aattttatat 60
tttttaaata tactatatat gttaaggata aggggtgaag ttttcttcct ttgtaatacc 120
tgttcaagag tttaatggat taggagatta gngttaacct tgaggataaa agtncaaatt 180
tgtctcatta ggacac 196

<210> 111
<211> 544
<212> DNA
<213> Homo sapien

<400> 111
gtcgacctca gccggtcgtc gcgacgttcg cccgctcgct ctgaggctcc tgaagccgaa 60
accagctaga ctttcctcct tcccgcctgc ctgtagcggc gttgttgcca ctccgccacc 120
atgttcgagg cgcgcctggt ccagggctcc atcctcaaga aggtgttgga ggcactcaag 180
gacctcatca acgaggcctg ctgggatatt agctccagcg gtgtaaacct gcagagcatg 240
gactcgtccc acgtctcttt ggtgcagctc accctgcggt ctgagggctt cgacacctac 300
cgctgcgacc gcaacctggc catgggcgtg aacctcacca gtatgtccaa aatactaaaa 360
tgcgccggca atgaagatat cattacacta agggccgaag ataacgcgga taccttggcg 420
ctagtatttg aagcaccaaa ccaggagaaa gtttcagact atgaaatgaa gttgatggat 480
ttagatgttg aacaacttgg aattccagaa caggagtact gctgtgtagt aaagatgcct 540
tctg 544

<210> 112
<211> 378
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(378)
<223> n = A,T,C or G

<400> 112
gtcgacacgg cttccgcacg gtcatccgcc ccttctacct gaccaactcc tcaggtgtgg 60
actagacggc gtggcccaag ggtggtgaga accggagaac cccaggacgc cctcactgca 120
ggctcccctc ctcggcttcc ttcctctctg caatgacctt caacaaccgg ccaccagatg 180
tcgccctact cacctgagcg ctcagcttca agaaattact ggaaggcttc cactagggtc 240
caccaggagt tctcccacca cctcaccagt ttccaggtgg taagcaccag gacgccctcg 300
aggttgctct gggatccccc cacagcccct ggncagtctg cccttgncac tggtctgaag 360
gtcattaaaa ttacattg 378

<210> 113
<211> 530
<212> DNA
<213> Homo sapien

<400> 113
gtcgacgtcg ttgtctttct aggtctcagc cggtcgtcgc gacgttcgcc cgctcgctct 60
gaggctcctg aagccgaaac cagctagact ttcctccttc ccgcctgcct gtagcggcgt 120
tgttgccact ccgccaccat gttcgaggcg cgcctggtcc agggctccat cctcaagaag 180

gtgttggagg cactcaagga cctcatcaac gaggcctgct gggatattag ctccagcggt 240
gtaaacctgc agagcatgga ctcgtcccac gtctctttgg tgcagctcac cctgcggtct 300
gagggcttcg acacctaccg ctgcgaccgc aacctggcca tgggcgtgaa cctcaccagt 360
atgtccaaaa tactaaaatg cgccggcaat gaagatatca ttacactaag ggccgaagat 420
aacgcggata ccttggcgct agtatttgaa gcaccaaacc aggagaaagt ttcagactat 480
gaaatgaagt tgatggattt agatgttgaa caacttggaa ttccagaaca 530

<210> 114
<211> 178
<212> DNA
<213> Homo sapien

<400> 114
gtcgacattt cttcctaata ttctataatc tccaactcct gaaaacccct ctctcaacta 60
atactttgct gttgaaatgt tgtgaaatgt taagtgtctg gaaatttttt ttttctaaga 120
aaaactatta aagtacttcc tagtagggca aaaaaaaaaa aaaaaaaaaa aaaaaaaa 178

<210> 115
<211> 211
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(211)
<223> n = A,T,C or G

<400> 115
ccttttttttt ttttttttng gntcaatctt ttatttggaa caaaggaaaa aaggactgac 60
accagtttag cctttgagtg tgcaaagctc tgccctccct cccacccctn agccccaaat 120
ccaanatttc atagccctaa cacccaccca agcagnttcc ctcacacatg ccctttgntt 180
tcttcctctc ttctatggtt ccttaggnaa a 211

<210> 116
<211> 439
<212> DNA
<213> Homo sapien

<400> 116
gtcgacctgt cactcactac atgaataagc aaatattgtc ttcaaaagaa tgcacaagaa 60
ccacaattaa gatgtcatat tattttgaaa gtacaaaata tactaaaaga gtgtgtgtgt 120
attcacgcag ttactcgctt ccatttttat gacctttcaa ctataggtaa taactcttag 180
agaaattaat ttaatattag aatttctatt atgaatcatg tgaaagcatg acattcgttc 240
acaatagcac tattttaaat aaattataag ctttaaggta cgaagtattt aatagatcta 300
atcaaatatg ttgattcatg gctataataa agcaggagca attataaaat cttcaatcaa 360
ttgaactttt acaaaaacca cttgagaatt tcatgagcac tttaaaatct gaactttcaa 420
agcttgctat taaatcatt 439

<210> 117
<211> 357
<212> DNA
<213> Homo sapien

<400> 117
gtcgactcca aattgacttt gcagcagggt ggcagggtca ggagagtctg gtcctgccta 60
gctcagattt catggcacct gcacttgaag caagtcactt ctttatcaca ggtgtcttga 120
aacattagct tcttttacca acctgagaaa attaggatga cctggcaaat aagatcttga 180
ataggccaaa agcaagtatc ttgctgtgtg tagtctcttg gttaaagtga agaaacagta 240

ctgttcacac ctttcttcac tgagattcca gtgtacatga gaacatatat ttattgcatg 300
attttctaga tacacagtct atgcattatt catatacatt tattttagcc taaagtg 357

<210> 118
<211> 431
<212> DNA
<213> Homo sapien

<400> 118
cctccctgag gaaattagga acctgttggc agatgttgaa acatttgtag cagatatact 60
gaaaggagaa aatttatcca agaaagcaaa ggaaaagaga gaatccctta ttaagaagat 120
aaaagatgta aagtctatct atcttcagga atttcaagac aaaggtgatg cagaagatgg 180
ggaagaatat gatgaccctt ttgctgggcc tccagacact atttcattag cctcagaacg 240
atatgataaa gacgatgaag ccccctctga tggagcccag tttcctccaa ttgcagcaca 300
agaccttcct tttgttctaa aggctggcta ccttgaaaaa cgcagaaaag atcacagctt 360
tctgggattt gaatggcaga aaacggtggt gtgctctcag taaaacggta ttctattatt 420
atggaagtga t 431

<210> 119
<211> 131
<212> DNA
<213> Homo sapien

<400> 119
cccctcgccc gtcacgcacc gcacgttcgt ggggaacctg gcgctaaacc attcgtagac 60
gacctgcttc tgggtcgggg tttcgtacgt agcagagcag ctccctcgct gcgatctatt 120
gaaaggtcga c 131

<210> 120
<211> 409
<212> DNA
<213> Homo sapien

<400> 120
gtcgacgtaa aagccacaca gaaatcaaaa gataagaata tagtttcagc taccaaaaag 60
cagcctcaga ataaaagtgc atttcagaag acaggaccca gctccttgaa gtctcctggc 120
cgtaccccac tgtccatcgt gagcctaccc cagtcttcta ccaaaacaca aactgcaccg 180
aagtcagcac agactgtcgc taagagccag cattcaacta aagggcctcc cagaagtggc 240
aaaaccccag cttcaatcag gaaaccaccc tcatctgtta aggatgcaga tagtggagat 300
aaaaaaccta ctgcaaagaa aaaggaagat gatgaccatt attttgtcat gactggaagt 360
aagaaaccta gaaaataaat acatactcat tataaaaaaa aaaaaaaag 409

<210> 121
<211> 131
<212> DNA
<213> Homo sapien

<400> 121
cccctcgccc gtcacgcacc gcacgttcgt ggggaacctg gcgctaaacc attcgtagac 60
gacctgcttc tgggtcgggg tttcgtacgt agcagagcag ctccctcgct gcgatctatt 120
gaaaggtcga c 131

<210> 122
<211> 130
<212> DNA
<213> Homo sapien

<400> 122

gtcgaccttt caatagatcg cagcgaggga gctgctctgc tacgtacgaa accccgaccc 60
agaagcaggt cgtctacgaa tggtttagcg ccaggttccc cacgaacgtg cggtgcgtga 120
cgggcgaggg 130

<210> 123
<211> 424
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(424)
<223> n = A,T,C or G

<400> 123
gtcgacgaga tgtggagtgg ctaaaagaag cctgtgttcc tgagaactta gaggaccagg 60
acctctattc caggcttgga cacctacatt tagactatta tatgaggaag caatcaactt 120
ctcacttgtt tcaaccactt tcacttgcag tcaaacctga attgtaagtg aaattgcttt 180
cctgatagca aacctgttgg attttctcca gaatccctgg gccactttta gcagtcagat 240
tcgtctaatc ctcctttaaa gatggtggca gtgaaactgg tacatgggac ctgactgggc 300
tttgtttgca actttctgat aatttataat tatttcaaaa taaaaaaatt ttaaaaataa 360
aaaaaaaaaa aaagggcggc cgctcggagt ctagagggcc cgtttaaacc cgntgatcag 420
cctc 424

<210> 124
<211> 548
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(548)
<223> n = A,T,C or G

<400> 124
ccttttttttt ttttttctc tagtaatgac tttattcatg aatctataat ggaattcaaa 60
atagcaaaga acatgaaaat gttcanatta atatttatta accaaatgca tcanaaaata 120
catctatttt cacatatcaa aagtgcctaa aatgcatgtg anaatataaa tattctccac 180
tttgnggaac ttcaagataa tgaaaaattg cttaatacac tttgccacaa aaactcatta 240
cactgcaaat ncagaanaaa taaaataact cattacattg cagatncaaa agaaatcaaa 300
tgtaactggc aaaataacca tttcatggct aatctttngg naaagngcta ttttcacact 360
gaaaaaaaga anttagaaaa gattaaaaat tttaaattct gaaccatcat tctnaaagtc 420
tgaagcgttt tctttagtat tcactatgtt catcacattc atgtgtnccc aacatgagac 480
taaacactat ctcaaaatct taaaaaatct ttccatncac anattatttc ctggaagnta 540
aaaattat 548

<210> 125
<211> 562
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(562)
<223> n = A,T,C or G

<400> 125
gtcgacgctc ctaacaaaga agatatcttg aaaatttcag aggatgagcg catggagctc 60

agtaagagct ttcgagtata ctgtattatc cttgtaaaac ccaaagatgt gagtctttgg 120
gctgcagtaa aggagacttg gaccaaacac tgtgacaaag cagagttctt cagttctgaa 180
aatgttaaag tgtttgagtc aattaatatg gacacaaatg acatgtggtt aatgatgaga 240
aaagcttaca aatacgcctt tgataagtat agagaccaat acaactggtt cttccttgca 300
cgccccacta cgtttgctat cattgaaaac ctaaagtatt ttttgttaaa aaaggatcca 360
tcacagcctt tctatctagg ccacactata aaatctggag accttgaata tgtgggtatg 420
gaaggaggaa ttgtcttaag tgtagaatca atgaaaagac ttaacagcct tctcaatatc 480
ccagaaaagt gtcctgaaca gggagggatg atttggaaga tatctgaaga taaacagcta 540
gcagnttgcc tgaaatatgc tg 562

<210> 126
<211> 131
<212> DNA
<213> Homo sapien

<400> 126
cccctcgccc gtcacgcacc gcacgttcgt ggggaacctg gcgctaaacc attcgtagac 60
gacctgcttc tgggtcgggg tttcgtacgt agcagagcag ctccctcgct gcgatctatt 120
gaaaggtcga c 131

<210> 127
<211> 512
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(512)
<223> n = A,T,C or G

<400> 127
gtcgacgtcc ggcttcggag cgggagtgtt cgttgtgcca gcgactaaaa agagaattaa 60
atatgggtga tgttgagaaa ggcaagaaga tttttattat gaagtgttcc cagtgccaca 120
ccgttgaaaa gggaggcaag cacaagactg ggccaaatct ccatggtctc tttgggcgga 180
agacaggtca ggcccctgga tactcttaca cagccgccaa taagaacaaa ggcatcatct 240
ggggagagga tacactgatg gagtatttgg agaatcccaa gaagtacatc cctggaacaa 300
aaatgatctt tgtcggcatt aagaagaagg aagaaagggc agacttaata gcttatctca 360
aaaaagctac taatgagtaa taattggcca ctgccttatt tattacaaaa cagaaatgtc 420
tcatgacttt tttatgtgta ccatccttta atagatctca tacaccagan tttcagatca 480
tgaatgactg acagaatatt ttgttgggca gt 512

<210> 128
<211> 483
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(483)
<223> n = A,T,C or G

<400> 128
gtcgacgttt ttgtgatact gacacatccc ccctttcaga acaccctctg cccttggatt 60
ctgtgcacag gaagctagtt gctcccctga atacactctt tcttccttgt aatacagcct 120
ctgattttga gcccaagaat aaagactaca gttctcagac tccttcgcaa ataaattttg 180
tgactaaact ctagtcaaca gtaagtgtca tgtagcagct cctgggaatc tcctttaaaa 240
agagagcttg tttataccta ttgtcatctc tgttcttctg tgccccttct tccattttgc 300
tgcctggaaa gcagatgtga tggctggaat tccagtcacc attttggacc atgaggacaa 360

caccctanag atgtggagtg gctaaaagaa gcctgtgttc ctgagaactt anaggaccan 420
gacctctatt ccaggcttgn acacctanat ttanactatt atatgaggaa gcaatcaact 480
tct 483

<210> 129
<211> 326
<212> DNA
<213> Homo sapien

<400> 129
gtcgaccttt tatctgtcta tccatccatc atcatttgaa ggcctaatat atgccaagta 60
ctcacatggt atgcattgag acataaaaaa gactgtctat aacctcaata agtattaaaa 120
atcccattat tacccataag gttcatctta tttcattttt agggaataaa attacatgtc 180
tatgaaattt caattttaag cactattgtt tttcatgacc ataatttatt tttaaaaata 240
aattaaaggt taattatatg catgtatgta tttctaataa ttaaaaatgt gttcaatccc 300
tgaaaaaaaa aaaaaaaaaa aaaaaa 326

<210> 130
<211> 276
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(276)
<223> n = A,T,C or G

<400> 130
gtcgacggac accagctgcg gaanttgcgg ctttggcaga ttgaaatcat ggcaggtcca 60
gaaagtgatg cgcaatacca gttcactggt attaaaaaat atttcaactc ttatactctc 120
acaggtagaa tgaactgtgt actggccaca tatggaagca ttgcattgat tgtcttatat 180
ttcaagttaa ggtccaaaaa aactccagct gtgaaagcaa cataaatgga ttttaaactg 240
tctacggttc ttaacctcat ctgttaagtt cccatg 276

<210> 131
<211> 482
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

<400> 131
ccttttttttt ttttttttaa attttaaggt tatttttatt tacaactttt gaaaaatgta 60
catttttttt tacatgggtt acttgtgcaa agttagattt ggaagtgata aatgcataaa 120
aggngacaat agaacattan acaaaacatt tacaagcctt gtcccatact gctacttaaa 180
ggtactatat atctaaaagt ataaatatcc aaaaaaagat cgcanacatt ggctttaagg 240
ttctcanatg ctgaaaggga anaaattaaa gcatgcagca ataactcagg atttgagtgg 300
aaaatagttn gccacanata tgctatgctc ccttccttga attcattaaa actctaaaat 360
aaagatggac aattgagttt attcacttag ggcagcactg atcctttaaa aagattaaag 420
gagctccaac tttccctagc tnaaaaactc acnatngttt ccattcctct gctcccacac 480
ct 482

<210> 132
<211> 428
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(428)
<223> n = A,T,C or G

<400> 132
ccttttttttt ttttttttgtc taaaaggcaa aaaactacaa acagcccaag tcctgagctc 60
cccaagacct ggatcctcca ctgtcccccct gaaacccggc aggaggcggg atggggagca 120
caanaggtgg gttcttaaaa aagtcacccc tggatgggaa agctcttcat cttctgccgc 180
cttcctntgc ctcccgctgc tgccgaggag agagatggan aggaccgggg ctatgccggc 240
aaactcaact tcttcccctt taggactttg gngatataga ggtanaanaa atcgcagtan 300
aggactgtct ggaccaggcc tgccacaatg gcnatgaggt cgaagaancc ctcgaaangg 360
taagcgccan anccagttga anagatanag cgtggcggta aacgcctagc gcaaacaagt 420
agnggctg 428

<210> 133
<211> 537
<212> DNA
<213> Homo sapien

<400> 133
gtcgacccca aacccactcc accttactac cagacaacct tagccaaacc atttacccaa 60
ataaagtata ggcgatagaa attgaaacct ggcgcaatag atatagtacc gcaagggaaa 120
gatgaaaaat tataaccaag cataatatag caaggactaa cccctatacc ttctgcataa 180
tgaattaact agaaataact ttgcaaggag agccaaagct aagacccccg aaaccagacg 240
agctacctaa gaacagctaa aagagcacac ccgtctatgt agcaaaatag tgggaagatt 300
tataggtaga ggcgacaaac ctaccgagcc tggtgatagc tggttgtcca agatagaatc 360
ttagttcaac tttaaatttg cccacagaac cctctaaatc cccttgtaaa tttaactgtt 420
agtccaaaga ggaacagctc tttggacact aggaaaaaac cttgtagaga gagtaaaaaa 480
tttaacaccc atagtaggcc taaaagcagc caccaattaa gaaagcgttc aagctca 537

<210> 134
<211> 535
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(535)
<223> n = A,T,C or G

<400> 134
gtcgactcct ctcacatggt ggctttagga agatccttgg ccaggagggt gatgccagct 60
atcttgcttc tgaaatatct acctgggatg gagtgatagt aacaccttca gaaaaggctt 120
atgagaagcc accagagaag aaggaaggag aggaagaaga ggagaataca gaagaaccac 180
ctcaaggaga ggaagaagaa agcatggaaa ctcaggagtg acattccctt cactcctttt 240
cctacccaag ggggaagact ggagcctaag ctgcctgcta ctgggcttta catggtgaca 300
gacatttccg tgggataggg aagatagcag gaagaaaagt aaactccata gaagtgtcat 360
tccactgggt tttgatattg gcttagctgc cagtctccca tttgtgacct atgccatcca 420
tctataatgg aggataccaa catttcttcc taatattcta taatctccaa ctcctgaaaa 480
acccctctct caactaatac tttgctgttg aaatgttgng aaatgttaag tgtct 535

<210> 135
<211> 114
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(114)
<223> n = A,T,C or G

<400> 135
gtcgacctca gcgtcattca gaannnggaa aagaatcaat gtaactcaag aaaggatgaa 60
aatacccttt cttcccatcc acgtgtttcc atctcaatcc tcacagggtc ctgg 114

<210> 136
<211> 354
<212> DNA
<213> Homo sapien

<400> 136
agaagcgaga tgacgaaggg aacgtcatcg tttggaaagc gtcgcaataa gacgcacacg 60
ttgtgccgcc gctgtggctc taaggcctac caccttcaga agtcgacctg tggcaaatgt 120
ggctaccctg ccaagcgcaa gagaaagtat aactggagtg ccaaggctaa aagacgaaat 180
accaccggaa ctggtcgaat gaggcaccta aaaattgtat accgcagatt caggcatgga 240
ttccgtgaag gaacaacacc taaacccaag agggcagctg ttgcagcatc cagttcatct 300
taagaatgtc aacgattagt catgcaataa atgttctggt tttaaaaaat aaaa 354

<210> 137
<211> 347
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(347)
<223> n = A,T,C or G

<400> 137
gtcgacggcg agattacgag gcgaggctcg cgcgcccgcc cccgccctgg cccccagtgc 60
ccacccggtc ggcccggcac agccatgatc aaggcgatcc taatcttcaa caaccacggg 120
aagccgcggc tctccaagtt ctaccagccc tacagtgaag atacacaaca gcaaatcatc 180
agggagactt tccatttggt atctaagaga gatgaaaatg tttgtaattt cctagaagga 240
ggattattaa ttggaggatc tgacaacaaa ctgatttata gacattatgc aacgttatat 300
tttgtcttct gtgnnggatt cttnanaaag tgaacttggc attttag 347

<210> 138
<211> 434
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(434)
<223> n = A,T,C or G

<400> 138
ccttttttttt ttttttggt taaatgactt actgtgtaat tttatttcat attacacaaa 60
tgttaatcaa atgctgagta gacatgcaga tgacaagcag tatatgacaa actctgaana 120
aatagttaca tgtagagttt ctcanatttt tagtgtatct aanaattaac tgaagagttt 180
gttaagaatg caggcttaaa ggccaatcca cagattataa tttcatacaa acaggatgga 240
gcctaanaac ctgtaaatta ttaaacaact gattaaaaat agagaggttt ctatgaagtt 300
aggnntgtcc ttatttctta tttgaactgg acaagtagaa ggataatagg taggaccaag 360

tgagcattat cagaatcaaa gtagaggcaa taacaagcca aggtgtttta ncctanctaa 420
agaagctcgt cgac 434

<210> 139
<211> 553
<212> DNA
<213> Homo sapien

<400> 139
gtcgacctga ctataacagt gcctactatg ttaacattag atgaacaagt gaattagagg 60
atttttaaat gtgtatccat cagtgtatgg acacactccc tctaacttct tcaaaaaaca 120
aaaattcctg gtagagctaa gtggttttta gaagtttggt tttggtaact gatttctacg 180
agataattga acacttttta aaatagttga tcattatgtc aaacagccct caacagtaaa 240
cttaaattag gtagaattat agtaagctgg aagagaaaat gttcccaaag agcattagtc 300
cctttctggc accttattac agatgaataa attgagactc acagaaatta aatgacttag 360
ccccagttat ccaactaact ccttaatgtg aggccatgat taggaatagg cttctagtat 420
tcagtcccat attattttga ctgtgtaata ccacgtgcca ctttgatttt aaagtcaaat 480
ctcggcttga actgtatggg gaaaaaaaaa atctccagct ggctctgctg aatccccaga 540
ggggccctcc act 553

<210> 140
<211> 450
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(450)
<223> n = A,T,C or G

<400> 140
gtcgacgccg gtgagttggg tgccggtgga gtcgtgttgg tcctcagaat ccccgcgtag 60
ccgctgcctc ctcctaccct cgccatgttt cttacccggt ctgagtacga caggggcgtg 120
aatacttttt ctcccgaagg aagattattt caagtggaat atgccattga ggctatcaag 180
cttggttcta cagccattgg gatccagaca tcagagggtg tgtgcctagc tgtggagaag 240
agaattactt ccccactgat ggagcccagc agcattgaga aaattgtaga gattgatgct 300
cacataggtt gtgccatgag tgggctaatt gctgatgcta agactttaat tgataaagcc 360
agagtggaga cacagaacca ctggttcacc tacaatgaga caatgaacag nggagagtgt 420
gacccaagct gngtccaatc tgnctttgca 450

<210> 141
<211> 140
<212> DNA
<213> Homo sapien

<400> 141
acacacccct ccctcacaca gggctcgacc gccgctggca gttccagggc taaggatttc 60
ctgcacttac ttgtggagaa ggagttcata gctgggctcc tggaggggag atagagcttc 120
tctttcgttc ccgggtcgac 140

<210> 142
<211> 591
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(591)

<223> n = A,T,C or G

<400> 142
gtcgacctgg acttgcagtg taaacagaga cgctgcaaat tgcttgtgga cggtgtaggc 60
cgctgcaggc caccatgaac cggcttccgg atgactacga cccctacgcg gttgaagagc 120
ctagcgacga ggagccggct ttgagcagct ctgaggatga agtggatgtg cttttacatg 180
gaactcctga ccaaaaacga aaactcatca gagaatgtct taccggagaa agtgaatcat 240
ctagtgaaga tgaatttgaa aaggagatgg aagctgaatt aaattctacc atgaaaacaa 300
tggaggacaa gttatcctct ctgggaactg gatcttcctc aggaaatgga aaagttgcaa 360
cagctccgac aaggtactac gatgatatat attttgattc tgattccgag gatgaagaca 420
gagcagtaca ggtgaccaag aaaaaaaaga agaaacaaca caagattcca acaaatgacg 480
aattactgta tgatcctgaa aaagataaca gagatcaggc ctgggttgat gcacagngaa 540
aggggttacc atggtttggg ancacaggag atcacgtcaa caacagcctg t 591

<210> 143
<211> 538
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(538)
<223> n = A,T,C or G

<400> 143
gtcgacaaat aagaagacac cttcagcatc ttaaactaga ataaataaaa gaagggtggc 60
ctcctagaat ttaagtcagg agggaggtgg tgggcaatgg atgacaagct ctactttgaa 120
gaggttgaat ttcagctgac cactactaaa gcagtacaag cttttccttt cagcaagtgt 180
cttcccagaa atgtgatagc aatttttagg aagaatttgg caaacataat gtttagcaga 240
tttgcaacaa atgctataag ctcaaatttt ttttttttt tttttnggca gcacactcag 300
ccctccaagg ggaagtggat tatttttctt gcaagtgcat tancanggga ggtattaagg 360
acagcaacat tccttcctgt ataaaaaaat aaataaataa aagaagaaag gattattgag 420
gccctctctg ctgnatgtaa tgtacttcan gatgttggta naaaagatat caacctanaa 480
taagnttcac aanaatacat ttggtttcac ngaaagttta aagtcaatct ggacattc 538

<210> 144
<211> 401
<212> DNA
<213> Homo sapien

<400> 144
gtcgacctgt tccctttttg ggcctgtctc cccatgtata tgttgagggg ttggacttca 60
gggcctgtga gaggccttcc aacttagact ttctccccag gagcataaat tcagtgaatc 120
tacgtgactc tcagtgatgg catcattgcc taatatccac ccagcttctg cttgaaaact 180
tccagagact ggttcacatg ggggtataaa agcccaggcc ccttgcccca acttgggaca 240
actatgaaga gtttccagct ccacagctcc ctgaggggct ggccgaggcc tttgtggggt 300
ttgcctcaca acccaattta tccctctggc caattctgct tcaatcactc cctgccaggt 360
gttgaccttg aatgtactcc cccaataaac ctcctgcaag c 401

<210> 145
<211> 367
<212> DNA
<213> Homo sapien

<400> 145
ccttttttct ttttttttag ttagaaatta caagtttatt tttatatttt gaaaaaggca 60
taatagaaaa caaaaataaa caaccaggca tatcaatatt tgtgacatac acatacacac 120
aaaaatgaat ataggaaata acacgaagaa aaagcatagt atgttttgaa accaacgtgg 180

ggcatgaaca gatttttgat gaaatacaac taaaggtttt aagtgtctat gtaatgttcg 240
agatattacg atcactctta tcctactagc aaaaattagc aaactaggct ttaaaacatg 300
attcctgttg ttttagcagg atttattttg gtaatgatcc tgcttcctta taaacaacta 360
cgtcgac 367

<210> 146
<211> 395
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(395)
<223> n = A,T,C or G

<400> 146
gtcgacaaga aagccccctt aatgttttta actgatgata ttttttaag cttaccaata 60
taagtatttt taaaggttct atttttcaaa gtcataacaa tgattgttct tgttttctct 120
catagaatag actgccatcg gataaagagt ggtccctagc ttctattttt ccaagtaaat 180
aagtagaaca tgttcttggg attataccat taaatgttaa ttttcttgaa gaagaaagat 240
tgttgtctgc caagatttta tgttagcgct cggattgagg cagaaaacgg aagcaccagg 300
tttaacactg ggatgacttg ggttgtgttc ctggaggttt gaagngggcc ttccccgcct 360
tttgaggggg aaaactgact gntttgaaca catat 395

<210> 147
<211> 455
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(455)
<223> n = A,T,C or G

<400> 147
gtcgactaaa aactggaacg gtgaaggtga cagcagtcgg ttggagcgag catcccccaa 60
agttcacaat gtggccgagg actttgattg cacattgttg ttttttttaat agtcattcca 120
aatatgagat gcgttgttac aggaagtccc ttgccatcct aaaagccacc ccacttctct 180
ctaaggagaa tggcccagtc ctctcccaag tccacacagg ggaggtgata gcattgcttt 240
cgtgtaaatt atgtaatgca aaatttttttt aatcttcgcc ttaatacttt tttattttgt 300
tttattttga atgatgagcc ttcgtgcccc cccttccccc ttttttgtcc cccaacttga 360
gatgtatgaa ggcttttggt ctccctggga gtgggtggan gcagccaggg cttacctgta 420
cactggactt gagaccagtt gaaataaaag tgcac 455

<210> 148
<211> 518
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(518)
<223> n = A,T,C or G

<400> 148
gtcgacctca cgccttcgcc gtagcatctt tcgcagcgga ccgaagagaa gaaaagtagg 60
ccagagccga actctcttcc tgccaagatg tctattggtg tgccgattaa agtactgcat 120
gaggccgagg gccacattgt gacatgtgag acgaacaccg gtgaggtata tcgggggaag 180

ctcattgaag cagaggacaa catgaactgc cagatgtcca acatcacagt cacatacaga 240
gatggccgag tggcacagct ggagcaggta tacatccgtg gcagcaaaat ccgctttctg 300
attttgcctg acatgctgaa gaacgcaccc atgttaaaga gcatgaaaaa taaaaaccaa 360
ggctcagggg ctggccgagg aaaagctgct attctcaagg cccaagtggc cgcaagagga 420
agaggacgtg gaatgggacg tggaaacatc tttcaaaagc gaagggataa ttttctaagt 480
tgaacagaac tttgtccttt tttctttcan gttatctg 518

<210> 149
<211> 442
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(442)
<223> n = A,T,C or G

<400> 149
ccttttttt ttttttttct tttcataaaa tttttacttt atgaattaaa tacattgaga 60
aacagngaaa atatatttac agtcatttga agngggcact actaacatat ttaatttaaa 120
aaaatctttg ctgtttcttt gcctgtttct ttcaaagaga attttaaata tgactttagc 180
ttttaaaaaa tacaatangg aaataattac attcttaata tgaaaacatt ttacaacgta 240
tcaccatggt caattaattc tgaatatcac ttaaaagttg atgttaaaat gtaaagngaa 300
tatttccttt cttgttanaa aatcaaaaag attatctcat taaaaacacc ttnggnccta 360
agacttatga tctgaanatg nccttttgaa aagnatcttc catggctaca actaaaaaan 420
acccggtaac acttgtgcac gg 442

<210> 150
<211> 341
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(341)
<223> n = A,T,C or G

<400> 150
gtnnacctat tattacccca tgatacagtt tagaaaacaa attcatgcac taagtaaatg 60
gaccaaatcg taagtcactg ccttttgctc cagagttggc tgctttgatt actcctacac 120
ttaactagtc aactttaaag aaaaaaattt ttttttctgt gaaggaaatt aagtgcctat 180
tttcanagag ctaaaagcaa tcaaggcatc tactgtgtta ttttcctatc catgtngact 240
catgtttaag gttgactagg aagacataat cattggctgc taataacaaa tngatttctt 300
ttnataaaaa atttaaaaga gtntntaatg ctttatttta t 341

<210> 151
<211> 459
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(459)
<223> n = A,T,C or G

<400> 151
gtcgaccagg tcttgaccct ggtcaacaag agaataggcc tttaccgtca ctttgacgag 60
accgtcaata ggtacaagca atcccgggac atctccaccc tcaacagtgg caagaagagc 120

ctggagactg aacacaaggc cttgaccagt gagattgcac tgctgcagtc caggctgaag 180
acagagggct ctgatctgtg cgacagagtg agcgaaatgc agaagctgga tgcacaggtc 240
aaggagctgg tgctgaagtc ggcggtggag gctgagcgcc tggtggctgg caagctcaag 300
aaagacacgt acattgagaa tgagaagctc atctcaggaa agcgccagga gctggtcacc 360
aagatcgacc acatcctgga tgccctgtag cccctgcccg catcctncag ggggcccagg 420
gtgcctgcac tttgctgtgg gnangcagat tgggtggta 459

<210> 152
<211> 242
<212> DNA
<213> Homo sapien

<400> 152
gtcgacccaa ggtcacagga gcattgcgtc gctgatgggg ttgaagtttg gtttggttct 60
tgtttcagcc caatatgtag agaacatttg aaacagtctg caccttttgat acggtattgc 120
atttccaaag ccaccaatcc attttgtgga ttttatgtgt ctgtggctta ataatcatag 180
taacaacaat aataccttttt tctccatttt gcttgcagga aacatacctt aagtttttttt 240
tg 242

<210> 153
<211> 57
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(57)
<223> n = A,T,C or G

<400> 153
ccttttttttt ttttttttttt ttccacatca ctcaggtttt atngaattta taaaatt 57

<210> 154
<211> 437
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(437)
<223> n = A,T,C or G

<400> 154
ccttttttttt ttttttttggt aatncagttt taatttattt tcatcacttt ttcttcataa 60
tccagatatt ttaaaatgca aagaaaatta actttcaatg atatgttcag ggactggcac 120
taaaaaaaat tttcagactg caaatgagtt atacaaatga aaatatcaaa tggagatcca 180
gttatcaaaa tgaaagcact caacatatta aaagttcaca agtatttgta ttgagcacat 240
tacaaaagtc agcttgctaa ctgttgtgat tttaaagaac tattgcanaa gtctgaanaa 300
aatanattta ttagttaact tataaagaga ttaaagaggc tgaaacaagt nttaaaaana 360
aatttgngcc tttattanaa tgttaggcgt cnacgcggcc gctcnngtct anagggcccg 420
tttaaacccg ctgatca 437

<210> 155
<211> 518
<212> DNA
<213> Homo sapien

<400> 155

gtcgacgtga gccacagtca cgccactgca ttctatcctg ggcaacagat ggagaccttg 60
tctcaaaaaa aaaaaattcc tgacatcgct atgtattccc aactttatca tttgtctgcc 120
tgtttagttt tgacttatgt ttttttttt tccccccctgt ggacatgtag ttgacggaaa 180
tcgtgaagga actttaatat tttatttaaa tttcccaaaa ctaatcatgc cttatgtgac 240
taatcttcag tgataatatt tcatctactg atatattttc ttgaggtgtg taattttcag 300
tataccttaa tcatttggta taaaaaagag agaggttttt gatatatgaa tgctgttctt 360
gtaaaaatca atcttgacac tttattttaa acttttttatt ggtaatgaca gtgggttttg 420
tacatcatga ttttcaattt aggatatctg tctaatttgt tttttcagag taactatatt 480
ggaattcaat aaaaatattc aaaatttttc ttaaaaaa 518

<210> 156
<211> 600
<212> DNA
<213> Homo sapien

<400> 156
gtcgacgttt atttaagttc atgtttcact gtttgcactt tgcattgaac aatgggttta 60
ttcgctgatg taaacggttc gagtgaagaa ttaatgcagt aagtatgaca acacatacac 120
acttgcctct ccccatctcc agaagagggg agcagagtcc gagcttatct aaatatgaat 180
gtggccacaa agctgtggaa ggtgacaaag cttaaacacc tttgccctgg ctctgcattg 240
tcacctagag agcaagaggt ctatagaaac atcatgtcac atgaaacgat tctctgcttt 300
ttggttctga acttgaagtc cctaaactgc aaaatctaag agttgggtgg ttattaaaat 360
gcttttaaaa agttaactgt ggcaccaatt ctaatgtaat ccaacttgtg actgtttttt 420
tttgttttgt tttgtttttg tgtgtgtgtg tgtgtggcac tgggaaaagt ggaaacaaac 480
atgtattgaa atacatattg gaaataaaaa tggtttgagc gtcagtgata ttctcccaga 540
atgtacttat cttacctcgc atgtactgta gtcactcagt atttgtatat gttgctagaa 600

<210> 157
<211> 542
<212> DNA
<213> Homo sapien

<400> 157
gtcgacggct gggaagtcag ttcgttctct cctctcctct cttcttgttt gaacatggtg 60
cggactaaag cagacagtgt tccaggcact tacagaaaag tggtggctgc tcgagccccc 120
agaaaggtgc ttggttcttc cacctctgcc actaattcga catcagtttc atcgaggaaa 180
gctgaaaata aatatgcagg agggaacccc gtttgcgtgc gcccaactcc caagtggcaa 240
aaaggaattg gagaattctt taggttgtcc cctaaagatt ctgaaaaaga gaatcagatt 300
cctgaagagg caggaagcag tggcttagga aaagcaaaga gaaaagcatg tcctttgcaa 360
cctgatcaca caaatgatga aaaagaatag aactttctca ttcatctttg aataacgtct 420
ccttgtttac cctggtattc tagaatgtaa atttacataa atgtgtttgt tccaattagc 480
tttgttgaac aggcatttaa ttaaaaaatt taggtttaaa tttagatgtt caaaagtagt 540
tg 542

<210> 158
<211> 526
<212> DNA
<213> Homo sapien

<400> 158
cacctcaggc tgtggctctt tgggcttctt cctaatgcag aagaagttgc ccagcagcaa 60
aatcagggag gaggtgagca cctcggcccc cgccaggatg aacacgtaca tgtagacgtg 120
ggtcgcatcc aggagtttgc ctcccgaagg gggcccgacg agcacggcca ccgcctccat 180
cagcagcacc aggccaatgg cactggagaa cttgtaggag atgccaaaga agatgcagaa 240
gaccacgagg ccgccgtagt cgcccgccgt agagcccgcc aggtccgcga ggccgttgaa 300
gaacatggag aagctgaaga ggtagacgga gtagggccgc accttcccaa gccccgccac 360
gaagcccgcg gccggccgcg cgaagatgtc aatgaagccc aggatggtga gcaggaaggc 420
ggccttggtg tcgggcacgc ccaggtcctt ggcgtagctc accacgaaca cgggcgggac 480

gaagagcccc agcaccatga ccgaggcggc cacggcgtaa agcaca 526

<210> 159
<211> 306
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(306)
<223> n = A,T,C or G

<400> 159
ccttttttt ttttttttt ttttttngga tgtatnngaa atttttttcta tatanatcat 60
gtgtgacttc cataaagaaa aataaacacc tatncacagt ttacctaata tgtgtaatgt 120
taatgaaaag aatcaaagaa agatgttcgt tcattaactc tctaaatnaa attgtttttc 180
catttttacc aacttgatac cttaatcaag ncactcttgt tcttccttaa gtgcaaatga 240
atttttgtt tgggttgggg gacaacacaa aatacaaacc tgggttggat tcactgaaag 300
gcccaa 306

<210> 160
<211> 528
<212> DNA
<213> Homo sapien

<400> 160
ctgaagagcg gcttgctctt cacatcctca ggactcaggg gctggtccct gagcacgtgg 60
aaacaaggac tttgcacagc accttccagc ccaacatttc ccagggaaaa cttcagatgt 120
gggtggatgt tttccccaag agtttggggc caccaggccc tcctttcaac atcacacccc 180
ggaaagccaa gaaatactac ctgcgtgtga tcatctggaa caccaaggac gttatcttgg 240
acgagaaaag catcacagga gaggaaatga gtgacatcta cgtcaaaggc tggattcctg 300
gcaatgaaga aaacaaacag aaaacagatg tccattacag atctttggat ggtgaaggga 360
attttaactg gcgatttgtt ttcccgtttg actaccttcc agccgaacaa ctctgtatcg 420
ttgcgaaaaa agagcatttc tggagtattg accaaacgga atttcgaatc ccacccaggc 480
tgatcattca gatatgggac aatgacaagt tttctctgga tgactact 528

<210> 161
<211> 527
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(527)
<223> n = A,T,C or G

<400> 161
ccttttttt ttttttttgg tcttacaact ctattgtaaa ctatactaga ctatagaggg 60
acttctacat ctttcaagat gtgtttaata aaggtctgtt tataataact tttgaggcat 120
gaatctagca aatagtactt tatacaatgt cccttgtcat taccaactca taaatattaa 180
gtgtttttca gtgacttatg tttggatgtg gtagtgctga tcagggccat gtgctgatgt 240
cctggagagc aaaatcaatc caaagnggng ctgctatttg tgacagaaca tgtttattta 300
ctcagccccg gagacaaaag gaaaattgat atgggggagc gggaaatagg agaactatta 360
aatgtagtga agaaatttca caggtctaaa ggaactatta aaaggaagga taaagtagat 420
tctatactat aaaacagaat cctacctctg ataaaagaca aatcagcctg aatttttgaa 480
taatcaatag gattcaaaat gactattttc aattgcaatc tcattct 527

<210> 162

<211> 77
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(77)
<223> n = A,T,C or G

<400> 162
cctttttttt ttttttttnt ttnnttttnt ttttttttnt ttttagggaa anaaatctgg 60
gttcctttta tttttga 77

<210> 163
<211> 645
<212> DNA
<213> Homo sapien

<400> 163
gtcgacaaac aatgaatagt ttttcattgt accatgaaat atccagaaca tacttatatg 60
taaagtatta tttatttgaa tctacaaaaa acaacaaata atttttaaat ataaggattt 120
tcctagatat tgcacgggag aatatacaaa tagcaaaatt gaggccaagg gccaagagaa 180
tatccgaact ttaatttcag gaattgaatg ggtttgctag aatgtgatat ttgaagcatc 240
acataaaaat gatgggacaa taaattttgc cataaagtca aatttagctg gaaatcctgg 300
atttttttct gttaaatctg gcaaccctag tctgctagcc aggatccaca agtccttgtt 360
ccactgtgcc ttggtttctc ctttatttct aagtggaaaa agtattagcc accatcttac 420
ctcacagtga tgttgtgagg acatgtggaa gcactttaag ttttttcatc ataacataaa 480
ttattttcaa gtgtaactta ttaacctatt tattatttat gtatttattt aagcatcaaa 540
tatttgtgca agaatttgga aaaatagaag atgaatcatt gattgaatag ttataaagat 600
gttatagtaa atttatttta ttttagatat taaatgatgt tttat 645

<210> 164
<211> 434
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(434)
<223> n = A,T,C or G

<400> 164
gtcgaccgga cgcggcggca ttaaacggtt gcaggcgtag cagagtggtc gttgtctttc 60
taggtctcag ccggtcgtcg cgacgttcgc ccgctcgctc tgaggctcct gaagccgaaa 120
ccagctagac tttcctcctt cccgcctgcc tgtagcggcg ttgttgccac tccgccacca 180
tgttcgaggc gcgcctggtc cagggctcca tcctcaagaa ggtgttggag gcactcaagg 240
acctcatcaa cgaggcctgc tgggatatta gctccagcgg tgtaaacctg cagagcatgg 300
actcgtccca cgtctctttg gtgcagctca ccctgcggtc tgagggcttn gacacctacc 360
gctgcgaccg caacctggcc atgggcgtga acctcaccag tatgtncaaa atactaaaat 420
gcgccngcaa tgaa 434

<210> 165
<211> 388
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

<222> (1)...(388)
<223> n = A,T,C or G

<400> 165
gtcgaccatt catatatata tgcatatata tgtgaagctc catatttctg ttgctttaaa 60
gaagtaaaac cttccattta aataagatga catgcntaan ataacaaagc ttccttgatt 120
tccttttcct gtgtaattna atagatttgt tgactagtgc ttgggcacat tataaatcag 180
ngttatttgc tcttggagcc atttttaaa aaaaattttg gcagtgagca gttgaattta 240
tcttgaattt atcatgtgtg tgtatttctg aagcagctac atagcagaac attttaagag 300
attctgttag cccacatgtt catgttggtt gctgctgaat ggtaaatatt aaataaaatt 360
accagattaa tcttaaaaaa aaaaaaaa 388

<210> 166
<211> 443
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(443)
<223> n = A,T,C or G

<400> 166
gtcgaccttg ctttcttaaa aaacaaaaaa actactgtca gtattaatac tgagccagac 60
tggcatctac agatttcaga tctatcattt tattgattct taagcttgta ttaaaaacta 120
ggcaatatca tcatggatac ataggagaag acacatttac aatcattcat tgggcctttt 180
atctgtctat ccatccatca tcatttgaag gcctaatata tgccaagtac tcacatggta 240
tgcattgaga cataaaaaag actgtctata acctcaataa gtattaaaaa tcccattatt 300
acccataagg ntcatcttat ttcattttta gggaataaaa ttacatgtct atgaaatttc 360
aattttaagc actattgntt ttcatgacca taatttattt ttaaaaataa attaaaggtt 420
aattataaaa aaaaaaaaaa aag 443

<210> 167
<211> 608
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(608)
<223> n = A,T,C or G

<400> 167
gtcgactgcg cctctccgaa cgcaacatga aggtgctcct tgccgccgcc ctcatcgcgg 60
ggtccgtctt cttcctgctg ctgccgggac cttctgcggc cgatgagaag aagaaggggc 120
ccaaagtcac cgtcaaggtg tattttgacc tacgaattgg agatgaagat gtaggccggg 180
tgatctttgg tctcttcgga aagactgttc caaaaacagt ggataatttt gtggccttag 240
ctacaggaga gaaaggattt ggctacaaaa acagcaaatt ccatcgtgta atcaaggact 300
tcatgatcca gggcggagac ttcaccaggg gagatggcac aggaggaaag agcatctacg 360
gtgagcgctt ccccgatgag aacttcaaac tgaagcacta cgggcctggc tgggtgagca 420
tggccaacgc aggcaaagac accaacggct cccagttctt catcacgaca gtcaagacag 480
cctggctaga tggcaagcat gtggtgtttg gcaaagttct agagggcatg gangtggtgc 540
ggaangtgga gagcaccaag acagacagcc gggataaacc cntgaangat gtgatcatcg 600
cagactgc 608

<210> 168
<211> 569
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(569)
<223> n = A,T,C or G

<400> 168
gtcgacgcgg ncggccggac agactgacgt gtgagctgca tcgcgggagg cgcatggngg 60
ggatggcgct ggcgcgggcc tggaagcaga tgtcctggtt ctactaccag tacctgctgg 120
tcacggcgct ctacatgctg gagccctggg agcggacggt gttcaattcc atgctggttt 180
ccattgtggg gatggcacta tacacaggat acgtcttcat gccccagcac atcatggcga 240
tattgcacta ctttgaaatc gtacaatgac caagatgcga ccaggatcag aggttncttg 300
gggaagaccc accctacgaa gttggaatga gaccatcaga tgtgataaga aactcttcta 360
gatgtcaaca taaccaacct tataaagact aaaattcatg agtagaacag gaaaatcatc 420
ctgactcatg tgttgtgttc tttatttta attttncaaa gaggctcttg tatagcagtt 480
ttttgtctat tttaacattg taagtcattt tgtnctttga natcantatt ttcttaacct 540
ttgtgactgt ttcaatatta cccccgnga 569

<210> 169
<211> 216
<212> DNA
<213> Homo sapien

<400> 169
gtcgaccggg aacccatcta taaagtaagg cacactcgta atggttgaat tgtgttctgg 60
ttaatttcct aaaggacttc acagttgcac ttatgaaaat gattttatat tgaaatgata 120
tttgcataag aaaaagcatg tgattaattg catattgctt gagtgttcat ctgtgaatgt 180
gaaaaataag ctgtttttt ttattagata tttgca 216

<210> 170
<211> 284
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(284)
<223> n = A,T,C or G

<400> 170
ccttttttt ttttttgaa atggancttc tgaatcgaaa agtttttcac tttaaatgtt 60
ggatgagtgc taccaaaaca ctnngcatct tagggcaagt gtcgctgagc acctgcttcc 120
ccatattctc agcannatca tttcagttct tagcaatctg gcaggcaaaa ggaaagtctg 180
attttgntng aattngcatt ttcctgatta ccancaaact antttaagct taatgggcac 240
ntnntatttc tattctctga actgcccatt tttctaccat tcag 284

<210> 171
<211> 541
<212> DNA
<213> Homo sapien

<400> 171
cagacagcac tgtgttggcg tacaggtctt tgcggatgtc cacgtcacac ttcatgatgg 60
agttgaaggt agtttcgtgg atgccacagg actccatgcc caggaaggaa ggctggaaga 120
gtgcctcagg gcagcggaac cgctcattgc caatggtgat gacctggccg tcaggcagct 180
cgtagctctt ctccagggag gagctggaag cagccgtggc catctcttgc tcgaagtcca 240
gggcgacgta gcacagcttc tccttaatgt cacgcacgat ttcccgctcg gccgtggtgg 300

tgaagctgta gccgcgctcg gtgaggatct tcatgaggta gtcagtcagg tcccggccag 360
ccaggtccag acgcaggatg gcatggggga gggcataccc ctcgtagatg ggcacagtgt 420
gggtgacccc gtcaccggag tccatcacga tgccagtggt acggccagag gcgtacaggg 480
atagcacagc ctggatagca acgtacatgg ctggggtgtt gaaggtctca aacatgatct 540
g 541

<210> 172
<211> 573
<212> DNA
<213> Homo sapien

<400> 172
gtcgactttc aacaaatcct gaagtctttc tgtgaagtga ccagttctga actttgaaga 60
taaataattg ctgtaaattc cttttgattt tctttttcca ggttcatggt ccttggtaat 120
ttcattcatg gaaaaaaatc ttattataat aacaacaaag atttgtatat ttttgacttt 180
atatttcctg agctctcctg actttgtgaa aaagggtgga tgaaaatgca ttccgaatct 240
gtgagggccc aaaacagaat ttaggggtgg gtgaaagcac ttgtgcttta gctttttcat 300
attaaatata tattatattt aaacattcat ggcatagatg atgatttaca gacaatttaa 360
aagttcaagt ctgtactgtt acagtttgag aattgtagat aacatcatac ataagtcatt 420
tagtaacagc ctttgtgaaa tgaacttgtt tactattgga gataaccaca cttaataaag 480
aagagacagt gaaagtacca tcataattaa cctaaatttt tgttatagca gagtttcttg 540
tttaaaaaaa aataaaatca tctgaaaagc aaa 573

<210> 173
<211> 545
<212> DNA
<213> Homo sapien

<400> 173
gtcgacctgg gctggacgtg gttttgtctg ctgcgcccgc tcttcgcgct ctcgtttcat 60
tttctgcagc gcgccagcag gatggcccac aagcagatct actactcgga caagtacttc 120
gacgaacact acgagtaccg gcatgttatg ttacccagag aactttccaa acaagtacct 180
aaaactcatc tgatgtctga agaggagtgg aggagacttg gtgtccaaca gagtctaggc 240
tgggttcatt acatgattca tgagccagaa ccacatattc ttctctttag acgacctctt 300
ccaaaagatc aacaaaaatg aagtttatct ggggatcgtc aaatcttttt caaatttaat 360
gtatatgtgt atataaggta gtattcagtg aatacttgag aaatgtacaa atctttcatc 420
catacctgtg catgagctgt attcttcaca gcaacagagc tcagttaaat gcaactgcaa 480
gtaggttact gtaagatgtt taagataaaa gttcttccag tcagtttttc tcttaagtgc 540
ctgtt 545

<210> 174
<211> 469
<212> DNA
<213> Homo sapien

<400> 174
gtcgacaaag aatcacagct ttctctccat gttttattaa cacacagaaa aatactttga 60
aaaatatacc atttctcaaa aatgaaatgt atgatttgct acaaatggcc atatggaaaa 120
tatgatacct gcttatttt gactcagggt gcattcaatt tttatactaa ctgaaaatta 180
catgattgcg ttttgttta aaagtgaaaa aaagtaataa ctgcttttag ccttgtaata 240
ttgaatgcgt caattggctc cccttgtaga atgttgaatg gctatcactg gtgacagatg 300
ttctgtacat cgcagtaata ctgcttatat aattgtgata attttccgct tcttatttgt 360
catttttagt gatttaaaaa tcccttgatg actccctgaa aaatgactga tgtttttcct 420
atattaagta atttctgctg gtaaagtgta agtcttttaa taatttctt 469

<210> 175
<211> 108
<212> DNA

<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(108)
<223> n = A,T,C or G

<400> 175
cctttttttt ttttttttng aaattnaagt aacttnatnn aaattcaaaa acaatnctta 60
aaactgnntt tagagtcaag acccttttgt attataaaaa tcacaagt 108

<210> 176
<211> 426
<212> DNA
<213> Homo sapien

<400> 176
gtcgactgtt tagaagttat acacagagag aaggggaaaa gaaactccat caatcaagct 60
aaaggcagca aaggaaaatt tgaaaagaag caacgagact gtttaacaaa gaacatcaaa 120
taagatgatg gaactagaag aaaaacacca atgtccttaa ttatataaaa acatcaatgt 180
ccttaattat ataaattttt aaccctcaat tgggttaaaa aatcagattt gtactaagag 240
atgtatcttt aaaagcaaaa gaaagaataa aaagatcaac aagtaaaaca aagtaggagt 300
cagaattaat attagacaaa ataaaggtga aaaatactaa atgcaagaaa taatatttta 360
gatgacaaaa atgtatgagc cataaaaaag tcatgagttt ttataaacct aaaatatagc 420
gtcgac 426

<210> 177
<211> 538
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(538)
<223> n = A,T,C or G

<400> 177
cctttttttt ttttttttt ttttttgga ngnattngaa atttttcta tatanatcat 60
gtgtgacttc cataaagaaa aataaacacc tatacacagt ttacctaata tgtgtaatgt 120
taatgaaaag aatcaaagaa agatgttcgt tcattaactc tntaaatcaa attgtttttc 180
catttttacc aacttgatac cttaatcaag tcactcttgt tcttccttaa gtgcaaatga 240
atttttgtt tgggttgggg gacaacacaa aatacaaacc tgggttggat tcactgaaag 300
gcccaanaaa gggccttant ctaggaagta nagngtgana tgatacaccc acaggctggn 360
gcattctggn ccacacaaan acgtgctgnt ccccgcccta ctgntnaaaa cagntctgtt 420
ttgctnanat gctgctgntg caacctgcag gtccatgana agaacaactc cctggttgtt 480
tacancccgn gagtgttttg ngaatttgca cctacatttc ccatgtgata tggactca 538

<210> 178
<211> 566
<212> DNA
<213> Homo sapien

<400> 178
gtcgacttgg aagcaggttt atttattata tacttgcaat tgaatataag atacagacat 60
atatatgtgt tatgtatttc tagaaatgca cataacatat atttgcctat tgtttaatgt 120
tttttccaga tatttattac agaagggcat ggagggatac ctacttattc ttcattatga 180
gaacaattaa aggcatttat tagataggaa attaacagat catctgcttc tataacttta 240
ttagctacat taaataggca gtgagcaata atttaaaaac tcaccattat ataaaataat 300

aaataacaaa gtaaaagtta atgttataaa aataaactga tagtaaggaa aatctaaatg 360
ggcatgatcc cattttagaa gaccaaatga ttaatagggt tgtcatgtta taatagacaa 420
ttgtctaatt atttctgtgt ttttatttag tgggtagcag aagttgttca gaagagcaga 480
aatatgtaga aaacatctct aaatttttgg caatttgaaa tagcaattct gaggcacaca 540
gctcatctac aaaaatcttt tgcaga 566

<210> 179
<211> 277
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(277)
<223> n = A,T,C or G

<400> 179
gncgacggga aaggaatatt atggcannaa gctgagcaag caattctggt ggaaagtcaa 60
acctgtcagt gctccacacc agggctgtgg tcctcccaga catgcatagg aatggccaca 120
ggtttacact gccttcccag caattataag cacaccagat tcagggagac tgaccaccaa 180
gggatagtgt aaaaggacat tttctcagtt gggtccatca gcagtttttc ttcctgcatt 240
tattgnngaa aactatngtt tcatttcttc ttttata 277

<210> 180
<211> 349
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(349)
<223> n = A,T,C or G

<400> 180
ccttttttttt ttttttttttt ttttttttttt ttttttttttt ttagnataag gaaaagctac 60
aaacctcaag gntgttttat ttaaaccaaa taatntgagc aagacatatn tacattaaaa 120
acaaatgaac acattaaaat ttcactattt tacaatctaa attctagcaa catatacaaa 180
tactgagnga ctacagtaca tgccgnggta ananaagtac attntgggan aatatnactg 240
acnctcaaac catttttatt tccaatatgt atttcaatac atgtttgttt ccacttttcc 300
cagngccaca cacacncnca cacaaaaaca aaacaaaaca aaaaaaaac 349

<210> 181
<211> 435
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(435)
<223> n = A,T,C or G

<400> 181
ccttttttttt ttttttttga catttacagg tatttatttg agtaagagct cataaaatat 60
atttttataa tatgcacaag aaaaaatаca tttgaatgaa taaaaaataa aatgacagga 120
ggtgacagaa tttagtgttt ataaatgagg tcataaagaa ctttaataat tcanagaana 180
agttcaaagt gtatttaaaa gttgagaccc tgctttacaa tattttataa ttttaaaaaa 240
aggcgtttaa aggtgatagg tgacttaata attttccact ttcaaaatgg gtttctagac 300
actgttatga agctgctatg tactaataat actttgcttg ccaaagtgtt tgggttttgt 360

tgttgtttgt ttgtttgttt gtttttggtt catgaacaac agtgtctaga aacccacttt 420
caaaatgggg tcgac 435

<210> 182
<211> 328
<212> DNA
<213> Homo sapien

<400> 182
gtcgaccatt gtatcttttt cttttctatc cctttacatt tactctttca gaatccttat 60
gttttactgt tttcagaaaa cttagttttt aaaatattct gctaatcatt tttcatataa 120
gtttacatta aataagtctt ttaaagttta ttataattaa ataaagttta ttttcacatg 180
tgttttcata tctactgtct cagaactttc tccttgccct atttttccta ttttatcccc 240
tttttgcatc ttttgagttg actttttatg attttatttt tctctcttta ctagtttgga 300
tattatctac cccactaata ttctttca 328

<210> 183
<211> 491
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(491)
<223> n = A,T,C or G

<400> 183
ccttttttnt ttttttnttt ttttttnttt ttacaaacct caaggttgtt ttatttaaac 60
caaataatct gagcaagaca tatatacatt aaaaacaaat gaacacatta aaatttcact 120
attttacaat ctaaattcta gcaacatata caaatactga gtgactacag tacatgccga 180
ggtaagataa gtacattctg gganaatatc actgacgctc aaaccatttt tatttccaat 240
atgtatttca atacatgttt gtttccactt ttcccagngc cacacacaca cacacaaaaa 300
caaaacaaaa caaaaaaaaa cagtcacaag ttggattaca ttanaattgg ngccacagtt 360
gactttaaaa gcattttaat aaccacccaa ctcttanatt ttgcagttta gggacttcaa 420
gttcanaacc aaaaagcana gaatcgtttc atgtgacatg atgtttctat agacctcttg 480
ctctctaggt c 491

<210> 184
<211> 478
<212> DNA
<213> Homo sapien

<400> 184
gtcgacggct gctgttggtt gggggccgtc ccgctcctaa ggcaggaaga tggtggccgc 60
aaagaagacg aaaaagtcgc tggagtcgat caactctagg ctccaactcg ttatgaaaag 120
tgggaagtac gtcctggggt acaagcagac tctgaagatg atcagacaag gcaaagcgaa 180
attggtcatt ctcgctaaca actgcccagc tttgaggaaa tctgaaatag agtactatgc 240
tatgttggct aaaactggtg tccatcacta cagtggcaat aatattgaac tgggcacagc 300
atgcggaaaa tactacagag tgtgcacact ggctatcatt gatccaggtg actctgacat 360
cattagaagc atgccagaac agactggtga aaagtaaacc ttttcaccta caaaatttca 420
cctgcaaacc ttaaacctgc aaaattttcc tttaataaaa tttgcttgtt ttaaaaaa 478

<210> 185
<211> 596
<212> DNA
<213> Homo sapien

<220>

<221> misc_feature
<222> (1)...(596)
<223> n = A,T,C or G

<400> 185
gtcgacggac gaggagtgcg gcactgatga gtactgcgct agtcccaccc gcggagggga 60
cgcgggcgtg caaatctgtc tcgcctgcag gaagcgccga aaacgctgca tgcgtcacgc 120
tatgtgctgc cccgggaatt actgcaaaaa tggaatatgt gtgtcttctg atcaaaatca 180
tttccgagga gaaattgagg aaaccatcac tgaaagcttt ggtaatgatc atagcacctt 240
ggatgggtat tccagaagaa ccaccttgtc ttcaaaaatg tatcacacca aaggacaaga 300
aggttctgtt tgtctccggt catcagactg tgcctcagga ttgtgttgtg ctagacactt 360
ctggtccaag atctgtaaac ctgtcctgaa agaaggtcaa gtgtgtacca agcataggag 420
aaaaggctct catggactag aaatattcca gcgttgttac tgtggagaag gtctgtcttg 480
ccggatacag aaagatcacc atcaagccag taattcttct aggcttcaca cttgncagag 540
acactaaacc agctatccaa atgcagtgaa ctccttttat ataatagatg ctatga 596

<210> 186
<211> 314
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(314)
<223> n = A,T,C or G

<400> 186
gtcgactgcc tatttaatgt agctaataaa gttatagaag cagatgatct gttaatttcc 60
tatctaataa atgcctttaa ttgttctcat aatgaagaat aagtaggtat ccctccatgc 120
ccttctgtaa taaatatctg gaaaaaacat taaacaatag gcaaatatat gttatgtgca 180
tttctagaaa tacataacac atatatatgt ctgtatctta tattcaattg caagtatata 240
ataaataaac ctgcttccaa acaacaaaaa aaaaaaaaaa aaaaaaaaan naaaaaaaaa 300
aaaaaaaaaa aaaa 314

<210> 187
<211> 331
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(331)
<223> n = A,T,C or G

<400> 187
ccttttttttt ttttttttatt cctcagngct tttgatttta attcttttgg catatctaaa 60
tgtcagaaag tgaatatata catacagaat tcaaaacacc ttcctaaaat ggttattatt 120
ggccantcat tnacatcttt attttgaaag tctgaattgn caaatagttc taaagtgcat 180
tcttgcagct aataaatagc agcatttgtt tataaaacct taagaaattc agaccagggc 240
tgganaagtc acaataaaaa atcagacatg atctanatat agtcttcctt aatcatctaa 300
gacaaacact tgtgtgaatt agtttataag g 331

<210> 188
<211> 567
<212> DNA
<213> Homo sapien

<400> 188

gtcgacgctg aagaaggaaa agaaatgtgt gaaactcata ggagttcccg ctgacgctga 60
ggccttaagt gaaagaagtg gaaacacccc taactctccc aggttagctg ctgaatcaaa 120
gcttcaaaca gaagttaaag aaggaaaaga aacttcaagc aaattggaaa aagaaacttg 180
taagaaatta caccctattc tatatgtgtc ttctaaatct actccagaga cccagtgccc 240
tcaacagtaa agacttgtct ttaataagag tacggtgcca cttgcctcaa aagttactat 300
ggtgcttaag attgtcttga tctgacatat atcaccttct gggttattta ctcattgtgc 360
caggacctgg cattttcatg tgcctttgac caagtgttca gaatttgctt gactctaacc 420
tggagagctt cttaagtgat gccccttcat ggagcttcta tgacagtgaa taaactatta 480
attgaaggaa aatgttataa ttaatgtatc tatttgctgc attgtatatg gattaaatga 540
taaaaaacaa gtaatctacc ctcagag 567

<210> 189
<211> 130
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(130)
<223> n = A,T,C or G

<400> 189
ccttttttt ttttttttt ttttttttt ttttttttt tttttatcnc ctaagnanat 60
tttaatataa attttgaaca gttataaaaa anaaanangg cctttgggtc aataacanaa 120
cataacaaaa 130

<210> 190
<211> 426
<212> DNA
<213> Homo sapien

<400> 190
gtcgaccaac ttcccacata tatttactaa gatgattaag acttacattt tctgcacagg 60
tctgcaaaaa caaaaattat aaactagtcc atccaagaac caaagtttgt ataaacaggt 120
tgctataagc ttggtgaaat gaaaatggaa catttcaatc aaacatttcc tatataacaa 180
ttattatatt tacaatttgg tttctgcaat atttttctta tgtccaccct tttaaaaatt 240
attatttgaa gtaatttatt tacaggaaat gttaatgaga tgtatttttct tatagagata 300
tttcttacag aaagctttgt agcagaatat atttgcagct attgactttg taatttagga 360
aaaatgtata ataagataaa atctattaaa tttttctcct ctaaaaactg aaaaaaaaaa 420
aaaaag 426

<210> 191
<211> 550
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(550)
<223> n = A,T,C or G

<400> 191
ccttttttttt ttttttttt ttttttttt tttagttngg gatatgacct ttattgaact 60
tatccaccan agnggaaata atgtctgtac aaaaccaaat gtttgttact ataacttctg 120
catcacaatt aaaatccaaa cagtttttta aaaacagtca actcaatcaa aacccactac 180
ttcanaatca atagcttntt tgaagccaca gtaacactta aatatggtta anactcgaat 240
gcanaaattt ggttggttgg aaagctaatt aaacttccaa cttgctcaaa tagaattaca 300
aaaaggcaaa attgtgtttt tcacananat acagnccact ggaatcacca acactggaca 360

gctgttanag tatttanagt cctganataa caaggaatcc aggcntcctt taaacagtct 420
tctgttgncc tttcttccca atcananatt tgtggatgtg tggaatgaca ccnccaccag 480
caattgtagc cttgatgann gaatccaatt cttcatctcc acgaatagca agttgcaagt 540
gacgaggggt 550

<210> 192
<211> 299
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(299)
<223> n = A,T,C or G

<400> 192
ccttttttt ttttttgaa attnnaaatt ttattacaaa aacttttat tgctataaga 60
aaaatatgta ttaattctac aaaataacat tcagattatg ttctaattca attattcaat 120
acaatttatt ctcttgtaaa taagagaaac ttatttagaa tataaaatta taacctaatg 180
acaaagctct agtaaattgn gaactacacc tctacaccgg gcttaaatgc atcctgatta 240
atgatttctt catacatgtc acttatttta tccaaaaaag gatttgagtt ctcgtcgac 299

<210> 193
<211> 536
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(536)
<223> n = A,T,C or G

<400> 193
ttttttttt ttttttttat tctnncaatt tttatttctc ttacatgctc aaagaagcca 60
agcaaatcca ggtatacatg tatatgtttt aattttacag gagagagaaa gaggtataag 120
gcaagaatta actacatttt catttcacta tttctttatg agctctattt tgctgctaag 180
ttcaagtttc aaaaaaatta ttaattcctc tgctatgtta tcttgtccca attcacaaaa 240
taacagggat ttccccatgt gactcaaaag caagaatctt actcctaaat aacataaaca 300
gcaatatgtg tgactactgt cattcattaa cttcgatggt gaagttcatt aaactgacca 360
ttaaaagaac atttgaacaa ttccaaaagg gagcaaggat aaatctccaa atcacccaat 420
agacaaggaa cccagagatg acatacagng tgctcacttc cacccactgc cactgagaac 480
actgattgct ctcttcaaac acagagcgaa gaatgggcct catgtcacat ggggca 536

<210> 194
<211> 566
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(566)
<223> n = A,T,C or G

<400> 194
gtcgactgca ctattaccca gggcagatat tatgagaaac tgtttcttct ctaagggttt 60
atggcagact ttgctttttt aacatgtgag aaatgaattt tttattttgt gatttatgtg 120
atttcttttg ctgagtgaag gaaaggagaa attgttgcta ttgtcagcat cttaaaggta 180

tttccagtca aggcaaggct aagtgctttg tgatagtatt aagcaagtca tgttttgaat 240
ggattacctg tagtgactca ttggaatgat ataattatac aagtaatgcc aaaaaccaag 300
tcaaagccta attaaccaaa gcactcattt aaaaatcatc atgtttggac ctatctggac 360
ctctcagcac tgtaaaatag ttttggtttt gtggcatatg aatagctgtt taacaaatca 420
aagttagctn tttgcttctc agctttttg ggcaatacaa gttaagttct taatggggag 480
acattatcat ggcatgactt aagggaacat tggtttgtga aggaaaaaca gattatctaa 540
agccatctct atgtttctgt tcagat 566

<210> 195
<211> 217
<212> DNA
<213> Homo sapien

<400> 195
gtcgacataa ataaatggaa gaaatatcat gttcatgggc ttcaaaagtc aacagtaaag 60
atgccatttt ttcctaaatt gatctacagg ttcagtgcaa ttccttccga atctcaccag 120
ggtttttggt agacataaac aagtttattc taaaatttgt atggaaaggc acaggtcctg 180
gaataactaa agcaacctta caaaaaaaaa aaaaaag 217

<210> 196
<211> 391
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(391)
<223> n = A,T,C or G

<400> 196
gtcgacggac agacttagga gttttgttta gagcagttaa catctgaagt gtctaatgca 60
ttaacttttg taaggtactg aatacttaat atgtgggaaa ccctttgcg tggtccttag 120
gcttacaatg tgcactgaat cgtttcatgt aagaatccaa agtggacacc attaacaggt 180
ctttgaaata tgcatgtact ttatattttc tatatttgta actttgcatg ttcttgtttt 240
gttatataaa aaaattgtaa atgtttaata tctgactgaa attaaacgag cgaagatgag 300
caccaaaaaa aaaaaaaaaa aaaaaaaaaa aaaannnaaa aaaaaaaann aaaaaaaaaa 360
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 391

<210> 197
<211> 445
<212> DNA
<213> Homo sapien

<400> 197
gtcgactgga tctttatgtc aatgtgtaca tagtacaagc ttttttactg gaattgaggt 60
ttaaaaccac acactgccct tttggtggtg tgcctgttgg gccaaaaatt gggtgataat 120
gtagtgtcac tttctcagct caatgcagtt tctacttttt cttatgggaa aatttttcat 180
aaaaccttttt tgcaccaaaa cccaggggtg ttttttgcaa tatccttgtt atcctcgtag 240
tgtgccaagt cagaggcttt ctcttgccct tttcctgctg tgttctcagg cctcccaagg 300
gctgtttgac tcaacagtct acatccttcg ttgtgttttg gagaatgtgg gggtgggggt 360
cagagttcaa ggtgtctgtt ccctttttcct gtgaactctt tctagtccct atttggggag 420
ggtggctgga aacagatttt tgctg 445

<210> 198
<211> 463
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(463)
<223> n = A,T,C or G

<400> 198
gtcgacgtca gtattaatac tgagccagac tggcatctac agatttcaga tctatcattt 60
tattgattct taagcttgta ttaaaaacta ggcaatatca tcatggatac ataggagaag 120
acacatttac aatcattcat tgggcctttt atctgtctat ccatccatca tcatttgaag 180
gcctaatata tgccaagtac tcacatggta tgcattgaga cataaaaaag actgtctata 240
acctcaataa gtattaaaaa tcccattatt acccataagg ttcatcttat ttcatttta 300
gggaataaaa ttacatgtct atgaaatttc aattttaagc actattgttt ttcatgacca 360
taatttattt ttaaaaataa attaaaggtt aattatatgc atgtatgtat ttctaataat 420
taaaaatgtg ttcaatccct ganaaaaaaa aaaaaaaaaa aaa 463

<210> 199
<211> 129
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(129)
<223> n = A,T,C or G

<400> 199
gtcgaccggc gggcagctgc agcttctgct gctgaggccg ggattgctac gactgggact 60
gaagactcag acgatgccct gctgaagatg accatcagcc ancaagagtt tggccgnact 120
gggcttcct 129

<210> 200
<211> 523
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(523)
<223> n = A,T,C or G

<400> 200
ccttttttttt ttttttttttt tttttnaaat ctttatttaa aagtccatgc taataatgng 60
tttacatttt tacagttaca ttatgataga aactgttgga ttttttaaat atctaaaaca 120
atggcccact gaanaaagga acaattaact ctttaattaa ttccttagga taaataccca 180
naaatttaac agctagggca gacttntaat acaataccga aagtccttcc aaaaaccaag 240
nggttgccaa cttatgtccc ttagcattat aacattcttg agccaatagt gtaaaaatac 300
gctgacaatt ttataggcaa acattactca aggtatctta ctttccactt attactaaag 360
taattaaccc ctaaacagat gctcctcaac agngggacta catcctggta aacctatcat 420
aagttgaaac tatcaagttg aaatgcattt agtaccctga taaacctatc ataaagttga 480
aaatttgtaa attgaaccag tgtaaatcag aggccatntt act 523

<210> 201
<211> 532
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature

<222> (1)...(532)
<223> n = A,T,C or G

<400> 201
ccttttttttt ttttttttaca cttgagctta gccaaaaggc tgagaagcga ttttttttta 60
aaagctgttc tttaccatgg tttaaacgct aaaatgcata gctataaaaa caaaacactg 120
agctaatctg attacatcca gcttttgcac tcaatagccc ttgaccctcc agtcataagc 180
aagcctgtca ttcgcccagc cctgctatac attctcatta tagtttcgtt tcaaatccag 240
tgttacagaa acaaaacacc aagccctcaa tcatgctatg cgtatcttta tgtgtgcatg 300
tcttatgtat gtttaaaata aacatttta aatgttttag gccaggcttg gnggctcatt 360
cagttttagt ttgctttttt tttgccattc tttgttattt tgngaataag taaaacattt 420
aaatacttaa gtcacatctg tataaaaagt atattcatag gaaggaattt aacaatttta 480
ataaaactta ttagcatatc aatgagtttc aagatacacc tgaaactaaa tt 532

<210> 202
<211> 114
<212> DNA
<213> Homo sapien

<400> 202
ctccttggtg tggtcttctc tgagtgaatg tcacaaggcc ggtgacagga gggggtggag 60
gtgaggggac aaagtagagg ccgagggtca gtgcctttgg agaaagtcca gaga 114

<210> 203
<211> 304
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(304)
<223> n = A,T,C or G

<400> 203
gtcgaccttt ttttcccaac ttcttgcttt ctattggatt gttagggatt tctgtttttc 60
actttatttc tctctgctta tttgaaagct atacagcatg gttttctttc tttagggatc 120
actcttccac tttacttttt aaagatggat aaatttttata catttaaaaa atttaatctg 180
tatttgtatc ttcttcctga gtggacctta gcatgttata aatgctcact gaataattct 240
cattgttaat tagagtttgg ttttattntt ttaaanncaa tgtacttact tattcttagn 300
gtaa 304

<210> 204
<211> 581
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(581)
<223> n = A,T,C or G

<400> 204
cngcgttgtg aggtgagcnn tttcagaagc gcgatcccag gacacgtcgg gaagcaagca 60
tccntttagc tgcttggaaa gaggaccaaa gacggctaaa anntcatttg gaaatatctc 120
taaatatttg ttaccatgta taagctgcta aagagaaatt gggcccaaca aaactaattg 180
aataattgag gcagatttgt gtgtatcatc aaattctatc cagaagttga agaatctgaa 240
tttaaagatt gtgtgcattt aataagagga tgacctttca gtttaatttc actatagaag 300

accatctgga aaatgaatta acacccatta gagatggagc tttgaccctg gattcctcaa 360
aagagctgtc agtctcagaa agtcaaaaag gagaagagag ggacagaaaa tgttctgcag 420
aacaatttga cttgcctcag gatcacttgt gggaacataa gtcaatggaa aatgcagctc 480
cctctcaaga cacagacagt ccactcagtg cagccagcag ttcaaggaac ttggagccac 540
atggaaaaca gccctccttg agagctgcca aagagcatgc t 581

<210> 205
<211> 409
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(409)
<223> n = A,T,C or G

<400> 205
gccctgaaga acagtgcctg gatgtggtga cccactggat ccaggaaggt gaagaagggc 60
gtccaaagga tgaccgccac ctccgtggct gtggctacct tcccggctgc ccgggctcca 120
atggtttcca caacaacgac accttccact tcctgaaatg ctgcaacacc accaaatgca 180
acgagggccc aatcctggag cttgaaaatc tgccgcagaa tggccgccag tgttacagct 240
gcaaggggaa cagcacccat ggatgctcct ctgaagagac tttcctcatt gactgccggg 300
gccccatgaa tcaatgtctg gtagccacgc gngcgacgtc acagagacnc ggaaaaacca 360
aagctatatn ggtaaagagg ctgtgcaacc cgctctcaat gtgccaaca 409

<210> 206
<211> 561
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(561)
<223> n = A,T,C or G

<400> 206
gtntcatggg aaaggacatg tctctcgaag aaaggttata aaccctgaga tatgagggtt 60
tttttgagac atccgagcct gtttcgttcc gggntgggan caggaataac cctgacttct 120
gagctttcat aaccccagga tcctccagaa aatttgcggc gcgctgaggg aaaaccttgc 180
tgaagctgta cattggaatg cgtttacagt cattgtaatg gaagcaaaat acatgaagga 240
aaaactgtta tttgtatccc tgcttattgc acctgacgac tagttgcaga tggttttgtt 300
tacctaagaa aacttgtgat ataaatgaaa aaaacacctg ttttcctaga gtcattggtt 360
acaaatatgc ttcgtctaag agctatttgt ccattctcct ggagagtgtt tcaatttcga 420
cccatcagtt gtgaaccact aattattcag atgaataagt gtacagatga ggagcaaatg 480
tttggtttta ttgaaagaaa caaagccata ctttcagaaa agcaagtggg atgtgcattt 540
gatatgcttt ggaagcttca a 561

<210> 207
<211> 461
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(461)
<223> n = A,T,C or G

<400> 207

ggtntttcca gccaatgtga cctttaaaac ctatgaaggt ntnatgcaca gttcgtgtca 60
acaggaaatg atggatgtca agcaattcat tgataaactc ctacctccaa ttgattgacg 120
tcactaagag gccttgtgta gaagtacacc agcatcattg tagtagagtg taaacctttt 180
cccatgccca gtcttcaaat ttctaatgtt ttgcagtgtt aaaatgtttt gcaaatacat 240
gccgataaca cagatcaaat aatatctcct catgagaaat ttatgatctt ttaagtttct 300
atacatgtat tcttataaga cgacccagga tctactatat tagaatagat gaagcaggta 360
gcttcttttt tctcaaatgt aattcagcaa aataatacag tactgccacc agatttttta 420
ttacatcatt tgaaaattag cagtatgctt aatgaaaatt t 461

<210> 208
<211> 296
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(296)
<223> n = A,T,C or G

<400> 208
gatgaacatc catccnaatt ncgaagagcc tatattatac cctcttcaag aatttgcatg 60
gcatcaatat ctacaggaga aaaaaaggga actcaaaaat gaaacctggg aatattcttc 120
ctctgtgatt tcttttgtta atggtcagtt tctgggtgat gcattggatc tgcagaaatg 180
ggcccacgag gtgtgggata tagttgacat taaaccctct gcactttatg acgcactcac 240
tgaggatttt tccgctaagt tcttaagaga caccaagcat gatttcgtgt ttttgg 296

<210> 209
<211> 282
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(282)
<223> n = A,T,C or G

<400> 209
gcataataaa tgctttgagc ttcttgacta tcatatacct aaagaaagtg catcagagaa 60
tnatattcct gacttttnnc tgactggcaa aaagcnagct ttatcttgtc ttataggatg 120
cttagtttgc cactncactt caaaccaatg ggacagtcnt anatggngng acagtgttna 180
ancncaccaa aaggntncnt ttccntgggg ccancnctgt cntnancctc nctaanctat 240
ttgnanaatt ttaancncnn gttaantaaa aaaaaaaaaa aa 282

<210> 210
<211> 1445
<212> DNA
<213> Homo sapiens

<400> 210
ggcgttgtga ggtgagcttt ttcagaagcg cgatcccagg acacgtcggg aagcaagcat 60
ccccagagct gcttggaaag aggaccaaag acgtctaaaa agtcatttgg aaatatctct 120
aaatatttgt taccatgtat aagctgctaa agagaaattg ggcccaacaa aactaattga 180
ataattgagg cagatttgtg tgtatcatca aattctatcc agaagttgaa gaatctgaat 240
ttaaagattg tgtgcattta ataagaggat gacctttcag tttaatttca ctatagaaga 300
ccatctggaa aatgaattaa cacccattag agatggagct ttgaccctgg attcctcaaa 360
agagctgtca gtctcagaaa gtcaaaaagg agaagagagg gacagaaaat gttctgcaga 420

acaatttgac ttgcctcagg atcacttgtg ggaacataag tcaatggaaa atgcagctcc 480
ctctcaagac acagacagtc cactcagtgc agccagcagt tcaaggaact tggagccaca 540
tggaaaacag ccctccttga gagctgccaa agagcatgct atgcctaaag atttaaagaa 600
gatgttagaa aataaagtca tagaaacatt accaggtttc cagcatgtta agttatcagt 660
agtgaaaacc atcttgttga aagagaactt ccctggagaa aacatagttt caaaaagctt 720
ttcttctcac tctgatctga ttacaggtgt ttatgaggga ggcttaaaaa tctgggaatg 780
tacctttgac ctcctggctt atttcacaaa ggccaaagtg aaatttgctg ggaaaaaagt 840
cttggatctt ggttgtggat caggtttact aggtataact gcattcaagg gagggtccaa 900
agaaattcac tttcaagatt ataacagtat ggtgattgat gaagtaacct tacctaatgt 960
agtagctaac tccactttgg aagatgaaga aaatgatgta aatgagccag atgtgaaaag 1020
atgcaggaaa ccaaaagtaa cacaactata taaatgccga ttttttttctg gtgagtggtc 1080
tgagttttgt aagcttgtac taagtagtga aaaacttttt gtaaaatatg atctcattct 1140
cacctcagaa accatttaca acccagatta ttatagtaat ttgcaccaga ctttccttag 1200
actgttaagt aaaaatggac gtgtactttt ggccagcaaa gcacattatt ttggtgtagg 1260
tggaggtgtt catctctttc agaagtttgt agaagaaaga gatgttttta agaccagaat 1320
actcaaaata attgatgaag gattgaagag gttcataatt gaaataactt ttaagtttcc 1380
tggttaatta acattcactg agtatccaaa atgaaataaa cagaaggacc aaaaaaaaaa 1440
aaaaa 1445

<210> 211
<211> 414
<212> DNA
<213> Homo sapiens

<400> 211
aaaaagggaa ggaaggagag acagataact ctcagtcatt taaaaaacta caataaaata 60
ttatgaatta tcaattagat caaagttcct cacagctata tttatatagg taaaaaaaaa 120
ttaaataggc taaatgccca aaaatttaag actggcaaaa tatacttggc taaatactgt 180
gcgtctctat taaataccat gtttcagaag aattattaat gacatgagaa tatgctcaaa 240
atacatattg atatgtgcaa atacatattg caaagtaaga ttatagaatg atcctagttc 300
aaaaatgtca catatatatg tatttaaaaa aaaaggcagt taagatttac aacaaaatgt 360
tagtggtggg accttctggt aggaatacag attttttttt attcagaagt tttt 414

<210> 212
<211> 720
<212> DNA
<213> Homo sapiens

<400> 212
gtcgacgtaa aatagaaaca gaaggggact ttatcaacct gattaacttt ctcaacatgt 60
taaccctaca gttaacatta taatcaatgg tgaatcattg agtactttcc ttctaagatc 120
agaaacagtt caaagtccac tctcaccatt tctattcaac attgtactgg aatcccagcc 180
agtgcagtaa taccaataat aaaaaattaa agtcataaag attgaaaagg atgaagtaaa 240
gctatttcaa ttctatttag aagtatttag aaaccccaaa gaatctacaa aaaactaata 300
gaaataagtg aatatatgaa ggtcttacta tacaagatca acatatcaaa agcagtggta 360
tttaagaaaa ggttggagac tatttataat aaacagtggt tgaattttgt taatgctttt 420
tctgtatttt ttgaaatgat cttattattt ttctctttgc taaaaatgtg agtaaccttg 480
agttgacttt ctgtgtaaat caaccttgtg tcccaggaaa aaactccaat tgatcatgat 540
gtgttatcct ttttatacat tgctgtattc aatatgctaa tatatttatt ttttgtgtct 600
atttcatgag ggatatcagt atgtaattgt tttttcttgt tatatctttg ttggttttat 660
taatcaacat tatgctaact tcatacaata tattggaaca tgctccctcc ttttattttc 720

<210> 213
<211> 1114
<212> DNA
<213> Homo sapiens

<400> 213

gctcctaaca aagaagatat cttgaaaatt tcagaggatg agcgcatgga gctcagtaag 60
agctttcgag tatactgtat tatccttgta aaacccaaag atgtgagtct ttgggctgca 120
gtaaaggaga cttggaccaa acactgtgac aaagcagagt tcttcagttc tgaaaatgtt 180
aaagtgtttg agtcaattaa tatggacaca aatgacatgt ggttaatgat gagaaaagct 240
tacaaatacg cctttgataa gtatagagac caatacaact ggttcttcct tgcacgcccc 300
actacgtttg ctatcattga aaacctaaag tatttttgt taaaaaagga tccatcacag 360
cctttctatc taggccacac tataaaatct ggagaccttg aatatgtggg tatggaagga 420
ggaattgtct taagtgtaga atcaatgaaa agacttaaca gccttctcaa tatcccagaa 480
aagtgtcctg aacagggagg gatgatttgg aagatatctg aagataaaca gctagcagtt 540
tgcctgaaat atgctggagt atttgcagaa aatgcagaag atgctgatgg aaaagatgta 600
tttaatacca aatctgttgg gctttctatt aaagaggcaa tgacttatca ccccaaccag 660
gtagtagaag gctgttgttc agatatggct gttactttta atggactgac tccaaatcag 720
atgcatgtga tgatgtatgg ggtataccgc cttagggcat ttgggcatat tttcaatgat 780
gcattggttt tcttacctcc aaatggttct gacaatgact gagaagtggt agaaaagcgt 840
gaatatgatc tttgtatagg acgtgtgttg tcattatttg tagtagtaac tacatatcca 900
atacagctgt atgtttcttt ttcttttcta atttggtggc actggtataa ccacacatta 960
aagtcagtag tacatttta aatgagggtg gtttttttct ttaaaacaca tgaacattgt 1020
aaatgtgttg gaaagaagtg ttttaagaat aataattttg caaataaact attaataaat 1080
attatatgtg ataaattcta aaaaaaaaaa aaaa 1114

<210> 214
<211> 1495
<212> DNA
<213> Homo sapiens

<400> 214
gtaacggatg gtgcgccaac gtgagaggaa acccgtgcgc ggctgcgctt tcctgtcccc 60
aagccgttct agacgcggat gaagtgcaaa acaaacttct ccatagagga gttgttgcaa 120
agttccagtt tataccaaac agtaatcaga ttccattgga agctaaagat tttgagagcc 180
ttttgtacta tatgcaacta acttgatttc aagcttggga acttttaaaa aaaacattaa 240
agcaaaatga aaaatgcttt ctgaaagcag ctccttttg aaaggtgtga tgcttggaag 300
ccattttctg tgctttgatc cactaatgct aaggacacat taggattggt catggaaata 360
gaatgcacca ccatgagcat catcacctac aagctcctaa caaagaagat atcttgaaaa 420
tttcagagga tgagcgcatg gagctcagta agagctttcg agtatactgt attatccttg 480
taaaacccaa agatgtgagt ctttgggctg cagtaaagga gacttggacc aaacactgtg 540
acaaagcaga gttcttcagt tctgaaaatg ttaaagagtt tgagtcaatt aatatggaca 600
caaatgacat gtggttaatg atgagaaaag cttacaaata cgcctttgat aagtatagag 660
accaatacaa ctggttcttc cttgcacgcc ccactacgtt tgctatcatt gaaaacctaa 720
agtatttttt gttaaaaaag gatccatcac agcctttcta tctaggccac actataaaat 780
ctggagacct tgaatatgtg ggtatggaag gaggaattgt cttaagtgta gaatcaatga 840
aaagacttaa cagccttctc aatatcccag aaaagtgtcc tgaacaggga gggatgattt 900
ggaagatatc cgaagataaa cagctagcag tttgcctgaa atatgctgga gtatttgcag 960
aaaatgcaga agatgctgat ggaaaagatg tatttaatac caaatctgtt gggctttcta 1020
ttaaagaggc aatgacttat caccccaacc aggtagtaga aggctgttgt tcagatatgg 1080
ctgttacttt taatggactg actccaaatc agatgcatgt gatgatgtat ggggtatacc 1140
gccttagggc atttgggcat attttcaatg atgcattggt tttcttacct ccaaatggtt 1200
ctgacaatga ctgagaagtg gtagaaaagc gtgaatatga tctttgtata ggacgtgtgt 1260
tgtcattatt tgtagtagta actacatatc caatacagct gtatgtttct ttttcttttc 1320
taatttggtg gcactggtat aaccacccat taaagtcagt agtacatttt taaatgaggg 1380
tggtttttt ctttaaaaca catgaacatt gtaaatgtgt tggaaaaaag tgttttaaga 1440
ataataattt tgcaaataaa ctattaataa atattatatg tgataaattc taacc 1495

<210> 215
<211> 838
<212> DNA
<213> Homo sapiens

<400> 215

ggctgggaag tcagttcgtt ctctcctctc ctctcttctt gtttgaacat ggtgcggact 60
aaagcagaca gtgttccagg cacttacaga aaagtggtgg ctgctcgagc ccccagaaag 120
gtgcttggtt cttccacctc tgccactaat tcgacatcag tttcatcgag gaaagctgaa 180
aataaatatg caggagggaa ccccgtttgc gtgcgcccaa ctcccaagtg gcaaaaagga 240
attggagaat tctttaggtt gtcccctaaa gattctgaaa aagagaatca gattcctgaa 300
gaggcaggaa gcagtggctt aggaaaagca aagagaaaag catgtccttt gcaacctgat 360
cacacaaatg atgaaaaaga atagaacttt ctcattcatc tttgaataac gtctccttgt 420
ttaccctggt attctagaat gtaaatttac ataaatgtgt ttgttccaat tagctttgtt 480
gaacaggcat ttaattaaaa aatttaggtt taaatttaga tgttcaaaag tagttgtgaa 540
atttgagaat ttgtaagact aattatggta acttagctta gtattcaata taatgcattg 600
tttggtttct tttaccaaat taagtgtcta gttcttgcta aaatcaagtc attgcattgt 660
gttctaatta caagtatgtt gtatttgaga tttgcttaga ttgttgtact gctgccattt 720
ttattggtgt ttgattattg gaatggtgcc atattgtcac tccttctact tgctttaaaa 780
agcagagtta gatttttgca cattaaaaaa ttcagtatta attaaaaaaa aaaaaaaa 838

<210> 216
<211> 938
<212> DNA
<213> Homo sapiens

<400> 216
cacctcaggc tgtggctctt tgggcttctt cctaatgcag aagaagttgc ccagcagcaa 60
aatcagggag gaggtgagca cctcggcccc cgccaggatg aacacgtaca tgtagacgtg 120
ggtcgcatcc aggagtttgc ctcccgaagg gggcccgacg agcacggcca ccgcctccat 180
cagcagcacc aggccaatgg cactggagaa cttgtaggag atgccaaaga agatgcagaa 240
gaccacgagg ccgccgtagt cgcccgccgt agagcccgcc aggtccgcga ggccgttgaa 300
gaacatggag aagctgaaga ggtagacgga gtagggccgc accttcccaa gccccgccac 360
gaagcccgcg gccggccgcg cgaagatgtc aatgaagccc aggatggtga gcaggaaggc 420
ggccttggtg tcgggcacgc ccaggtcctt ggcgtagctc accacgaaca cgggcgggac 480
gaagagcccc agcaccatga ccgaggcggc cacggcgtaa agcacaaagc cgcggtcccg 540
gaagacgctc aggtctagca ggcgccggga gggtcgcggc ggccccgagc ccggctgggc 600
cgtgaccacc aggggcctca tgagtgcggc acacacgcag cagttgagca gcaggccgcc 660
caggatgagg aagccgcccc gccagccgta gcggtcctgc agcagctgcc ccagcgggct 720
cagggcacac aggaagacag ggctacctgc tgccgccagc ccgttggcca tggggcgccg 780
cttgctgaag tagcggttca gcatgatgag cgagggctgg aagttgagtg ccaaacccaa 840
ccccgtgatg accccagtgg tgaggtagac ctggatgatg ctccggcaaa aggacgcagc 900
caccatgccc agcgacgcaa agagaccccc cacaagca 938

<210> 217
<211> 1982
<212> DNA
<213> Homo sapiens

<400> 217
ggcgagaggc gggctgaggc ggcccagcgg cggcaggtga ggcggaacca accctcctgg 60
ccatgggagg ggccgtggtg gacgagggcc ccacaggcgt caaggcccct gacggcggct 120
ggggctgggc cgtgctcttc ggctgtttcg tcatcactgg cttctcctac gccttcccca 180
aggccgtcag tgtcttcttc aaggagctca tacaggagtt tgggatcggc tacagcgaca 240
cagcctggat ctcctccatc ctgctggcca tgctctacgg gacaggtccg ctctgcagtg 300
tgtgcgtgaa ccgctttggc tgccggcccg tcatgcttgt ggggggtctc tttgcgtcgc 360
tgggcatggt ggctgcgtcc ttttgccgga gcatcatcca ggtctacctc accactgggg 420
tcatcacggg gttgggtttg gcactcaact tccagccctc gctcatcatg ctgaaccgct 480
acttcagcaa gcggcgcccc atggccaacg ggctggcggc agcaggtagc cctgtcttcc 540
tgtgtgccct gagcccgctg gggcagctgc tgcaggaccg ctacggctgg cggggcggct 600
tcctcatcct gggcggcctg ctgctcaact gctgcgtgtg tgccgcactc atgaggcccc 660
tggtggtcac ggcccagccg ggctcggggc cgccgcgacc ctcccggcgc ctgctagacc 720
tgagcgtctt ccgggaccgc ggctttgtgc tttacgccgt ggccgcctcg gtcatggtgc 780
tggggctctt cgtcccgccc gtgttcgtgg tgagctacgc caaggacctg ggcgtgcccg 840

acaccaaggc cgccttcctg ctcaccatcc tgggcttcat tgacatcttc gcgcggccgg 900
ccgcgggctt cgtggcgggg cttgggaagg tgcggcccta ctccgtctac ctcttcagct 960
tctccatgtt cttcaacggc ctcgcggacc tggcgggctc tacggcgggc gactacggcg 1020
gcctcgtggt cttctgcatc ttctttggca tctcctacgg catggtgggg gccctgcagt 1080
tcgaggtgct catggccatc gtgggcaccc acaagttctc cagtgccatt ggcctggtgc 1140
tgctgatgga ggcggtggcc gtgctcgtcg ggcccccttc gggaggcaaa ctcctggatg 1200
cgacccacgt ctacatgtac gtgttcatcc tggcgggggc cgaggtgctc acctcctccc 1260
tgattttgct gctgggcaac ttcttctgca ttaggaagaa gcccaaagag ccacagcctg 1320
aggtggcggc cgcggaggag gagaagctcc acaagcctcc tgcagactcg ggggtggact 1380
tgcgggaggt ggagcatttc ctgaaggctg agcctgagaa aaacggggag gtggttcaca 1440
ccccggaaac aagtgtctga gtggctgggc ggggccggca ggcacaggga ggaggtacag 1500
aagccggcaa cgcttgctat ttattttaca aactggactg gctcaggcag ggccacggct 1560
gggctccagc tgccggccca gcggatcgtc gcccgatcag tgttttgagg gggaaggtgg 1620
cggggtggga accgtgtcat tccagagtgg atctgcggtg aagccaagcc gcaaggttac 1680
aaggcatcct caccaggggc cccgcctgct gctcccaggt ggcctgcggc cactgctatg 1740
ctcaaggacc tggaaaccca tgcttcgaga caacgtgact ttaatgggag ggtgggtggg 1800
ccgcagacag gctggcaggg caggtgctgc gtggggccct ctccagcccg tcctaccctg 1860
ggctcacatg gggcctgtgc ccacccctct tgagtgtctt ggggacagct ctttccaccc 1920
ctggaagatg gaaataaacc tgcgtgtggg tggagtgttc tcgtgccgaa ttcaaaaagc 1980
tt 1982

<210> 218
<211> 592
<212> DNA
<213> Homo sapiens

<400> 218
aggtctcatg ggaaaggtca tgtctctcga agaaaggtta taaaccctga gatatgaggg 60
ttgggcgaga catccgagcc tgtttcgttc cgtgttggga ccaggaataa ccctgacttc 120
tgagctttca taaccccagg atcctccaga aaatttgcgg cgcgctgagg gaaaaccttg 180
ctgaagctgt acattggaat gcgtttacag tcattgtaat ggaagcaaaa tacatgaagg 240
aaaaactgtt atttgtatcc ctgcttattg cacctgacga ctagttgcag atggttttgt 300
ttacctaaga aaacttgtga tataaatgaa aaaaacacct gttttcctag agtcattggt 360
tacaaatatg cttcgtctaa gagctatttg tccattctcc tggagagtgt ttcaatttcg 420
acccatcagt tgtgaaccac taattattca gatgaataag tgtacagatg aggagcaaat 480
gtttggtttt attgaaagaa acaaagccat actttcagaa aagcaagtgg gatgtgcatt 540
tgatatgctt tggaagcttc aaaagcagaa gaccagcctg ttaaaaaatg ct 592

<210> 219
<211> 650
<212> DNA
<213> Homo sapiens

<400> 219
ctgctaccca tccctttatg aagaggtttt gggagaggag caagagggag tctgagcacc 60
agccgcagcc ggggccaaag tttgtggggt cagggcccca tccagcagct gccctgcccc 120
atgtgacatg aggcccattc ttcgctctgt gtttgaagag agcaatcagt gttctcagtg 180
gcagtgggtg gaagtgagca cactgtatgt catctctggg ttccttgtct attgggtgat 240
ttggagattt atccttgctc ccttttggaa ttgttcaaat gttcttttaa tggtcagttt 300
aatgaacttc accatcgaag ttaatgaatg acagtagtca cacatattgc tgtttatgtt 360
atttaggagt aagattcttg cttttgagtc acatggggaa atccctgtta ttttgtgaat 420
tgggacaaga taacatagca gaggaattaa taatttttt gaaacttgaa cttagcagca 480
aaatagagct cataaagaaa tagtgaaatg aaaatgtagt taattcttgc cttatacctc 540
tttctctctc ctgtaaaatt aaaacatata catgtatacc tggatttgct tggcttcttt 600
gagcatgtaa gagaaataaa aattgaaaga ataaaaaaaa aaaaaaaaaa 650

<210> 220
<211> 782

<212> DNA
<213> Homo sapiens

<400> 220
ggtgaatcca gccaatgtga cctttaaaac ctatgaaggt atgatgcaca gttcgtgtca 60
acaggaaatg atggatgtca agcaattcat tgataaactc ctacctccaa ttgattgacg 120
tcactaagag gccttgtgta gaagtacacc agcatcattg tagtagagtg taaacctttt 180
cccatgccca gtcttcaaat ttctaatgtt ttgcagtgtt aaaatgtttt gcaaatacat 240
gccgataaca cagatcaaat aatatctcct catgagaaat ttatgatctt ttaagtttct 300
atacatgtat tcttataaga cgacccagga tctactatat tagaatagat gaagcaggta 360
gcttcttttt tctcaaatgt aattcagcaa aataatacag tactgccacc agatttttta 420
ttacatcatt tgaaaattag cagtatgctt aatgaaaatt tgttcaggta taaatgagca 480
gttaagatat aaacaattta tgcatgctgt gacttagtct atggatttat tccaaaattg 540
cttagtcacc atgcagtgtc tgtattttta tatatgtgtt catatataca taatgattat 600
aatacataat aagaatgagg tggtattaca ttattcctaa taatagggat aatgctgttt 660
attgtcaaga aaaagtaaaa tcgttctctt caattaatgg ccctttttatt ttgggaccag 720
gcttttattc tccctgatat tatttctatt taatactctt ttctctcaaa aaaaaaaaaa 780
aa 782

<210> 221
<211> 2417
<212> DNA
<213> Homo sapiens

<400> 221
cttccttccg cttgcgctgt gagctgaggc ggtgtatgtg cggcaataac atgtcaaccc 60
cgctgcccgc catcgtgccc gccgcccgga aggccaccgc tgcggtgatt ttcctgcatg 120
gattgggaga tactgggcac ggatgggcag aagcctttgc aggtatcaga agttcacata 180
tcaaatatat ctgcccgcat gcgcctgtta ggcctgttac attaaatatg aacgtggcta 240
tgccttcatg gtttgatatt attgggcttt caccagattc acaggaggat gaatctggga 300
ttaaacaggc agcagaaaat ataaaagctt tgattgatca agaagtgaag aatggcattc 360
cttctaacag aattattttg ggagggtttt ctcagggagg agctttatct ttatatactg 420
cccttaccac acagcagaaa ctggcaggtg tcactgcact cagttgctgg cttccacttc 480
gggcttcctt tccacagggt cctatcggtg gtgctaatag agatatttct attctccagt 540
gccacgggga ttgtgaccct ttggttcccc tgatgtttgg ttctcttacg gtggaaaaac 600
taaaaacatt ggtgaatcca gccaatgtga cctttaaaac ctatgaaggt atgatgcaca 660
gttcgtgtca acaggaaatg atggatgtca agcaattcat tgataaactc ctacctccaa 720
ttgattgacg tcactaagag gccttgtgta gaagtacacc agcatcattg tagtagagtg 780
taaacctttt cccatgccca gtcttcaaat ttctaatgtt ttgcagtgtt aaaatgtttt 840
gcaaatacat gccgataaca cagatcaaat aatatctcct catgagaaat ttatgatctt 900
ttaagtttct atacatgtat tcttataaga cgacccagga tctactatat tagaatagat 960
gaagcaggta gcttcttttt tctcaaatgt aattcagcaa aataatacag tactgccacc 1020
agatttttta ttacatcatt tgaaaattag cagtatgctt aatgaaaatt tgttcaggta 1080
taaatgagca gttaagatat aaacaattta tgcatgctgt gacttagtct atggatttat 1140
tccaaaattg cttagtcacc atgcagtgtc tgtattttta tatatgtgtt catatataca 1200
taatgattat aatacataat aagaatgagg tggtattaca ttattcctaa taatagggat 1260
aatgctgttt attgtcaaga aaaagtaaaa tcgttctctt caattaatgg ccctttttatt 1320
ttgggaccag gcttttattt tccctgatat tatttctatt taatactctt ttctctcaag 1380
aaaaaaaaaa aagtttgttt tttctttatt gtccttcata gcaggccaag tattgcctct 1440
ctgcaataga cagctactgt caatacatgc tgtaatttga cattctgggt cacagatata 1500
aggtatttaa aatctattta tgctttatag agaaaccaga cattaaaact tcatgcacta 1560
cttatttcga attactgtac cttatccaaa tttacaccta gctattagga tcttcaaccc 1620
aggtaacagg aataattctg tggtttcatt tttctgtaaa caactgaaag aataattaga 1680
tcatattcta gtatgttctg aaatatcttt aagactgatc ttaaaaacta acttctaaga 1740
tgatttcatc ttctcatagt atagagttta ctttgtacac gttgaaacca actactgtag 1800
aagatgagga atctattgta atttttttgct ttattttcat ctgccagtgg acttatttga 1860
attttcactt tagtcaaatt atttttttgta ttagtttttg atgcagacat aaaaatagca 1920
atcattttaa attgtcaaaa tttccagatt actggtaaaa attatttgaa aacaaactta 1980

tgggtaataa aggctagtca gaaccctata ccataaagtg tagttaccat acagattaat 2040
atgtagcaaa aatgtatgct tgatatttct caactgtgtt aatttttctg ctgtattcca 2100
gctgaccaaa acaatattaa gaatgcatct ttataaatgg gtgctaattg ataatggaaa 2160
taatttagta atggactata caggatgtta ataatgaagc catatgttta tgtctggatt 2220
taaaaatttt aaacaatcat ttactatgtc atttttcttt accttgaaga acataaactg 2280
ttatttcact tctacaaatc agcaagatat tatttatggc aagaaatatt ccattgaaat 2340
attgtgctgt aacatgggaa agtgtaaatg tttttcatgg tttctatcaa tgtgaaataa 2400
aatttaattc tgaaaaa 2417

<210> 222
<211> 1466
<212> DNA
<213> Homo sapiens

<400> 222
ggtggtgggg ctcttcagct gccccaactt tcagattgcg aagagcgccg ctgagaatct 60
gaagaataat catccatcca aatttgaaga tcctatatta gttcctcttc aagaatttgc 120
atggcatcaa tatctacagg agaaaaaaag ggaactcaaa aatgaaacct gggaatattc 180
ttcctctgtg atttcttttg ttaatggtca gtttctgggt gatgcattgg atctgcagaa 240
atgggcccac gaggtgtggg atatagttga cattaaaccc tctgcacttt atgacgcact 300
cactgaggat ttttccgcta agttcttaag agacaccaag catgatttcg tgtttttgga 360
catttgtatt gattcttctc caattggaag attgattttt gagctatact gtgatgtgtg 420
tcccaaaaca tgtaaaaatt ttcaggtctt gtgcacagga aaagcagggt tttctcaacg 480
tggcataaga ctacattaca aaaattccat ttttcatcga atagtacaga atggctggat 540
acaaggaggg gatatagtgt atggaaaagg agataatgga gagtcgattt atggtccaac 600
atttgaagga caggctccca tgcagatgga actgttggga atggcatcaa aattatccag 660
gatacaggga caatcatgac aaagagggaa taggaacaga gtcagaaatt taaggaagaa 720
agccacatgc ttcaatatgc aagattttca acatgcaaga gggagctttt tgaaactaga 780
aaatctactt tctttctaaa gacacatctt ctaaacattt aggaaaacta atgtcaccct 840
atataacaaa gagagtttct ctgaaagaaa ataatgttta ttcaggaata gggtattgct 900
gtaggcatac atgtgccata ggaaacgatg tgcatattca ggaaggtaaa ggcagacaaa 960
gggttttaaa ggaaaattgg ggaagattac gtaattgttt tgaaatgatt atccttggct 1020
atagcgatca gtaacaagag tccaaggttg gactggacag gtgtccctgc agaagtatta 1080
atatttcctg cataaggtcg caatggcctt tatgcaaggt tgtggctttt gtagttcttt 1140
gtgattgttt tgctatcagg catacaagtg tgagagttct ctgttcatag ctttccttgg 1200
ctctatttgt cagcattttt taaacatgac tacattttga ttctgaccac tattacacta 1260
attttatatt agaatgaaca atagaagttt caaggtgatt ataagaataa agagaataaa 1320
gagcagagta acatcagcac tgatagtgaa tgtaccctag aaagacatgc tcataggata 1380
cagttgaccc ttgagcaaca tggatttgaa atgtacgagt ccacttaaag aaacttacag 1440
tcagtttctt taaaaaaaaa aaaaaa 1466